Age-of-Information Revisited: Two-way Delay and Distribution-oblivious Online Algorithm

Cho-Hsin Tsai and Chih-Chun Wang
School of Electrical and Computer Engineering, Purdue University, USA
Email: {chtsai, chihw}@purdue.edu

***Abstract*—The ever-increasing needs of supporting real-time applications have spurred a considerable number of studies on minimizing Age-of-Information (AoI), a new metric characterizing the data freshness of the system. This work revisits and significantly strengthens the seminal results of Sun *et al.* on the following fronts: (i) The optimal waiting policy is generalized from the 1-way delay to the *2-way delay* setting; (ii) A new way of computing the optimal policy with *quadratic convergence rate*, an order-of-magnitude improvement over the state-of-the-art bisection methods; and (iii) A new low-complexity adaptive online algorithm that provably converges to the optimal policy without knowing the exact delay distribution, a sharp departure from the existing AoI algorithms. Contribution (iii) is especially important in practice since the delay distribution can sometimes be hard to know in advance and may change over time. Simulation results in various settings are consistent with the theoretic findings.**

I. Introduction

Thanks to the accelerating growth of networked systems in the past decades, the capability of providing real-time status updates has been the cornerstone of many important practical systems. Examples include remote health monitoring, GPS location tracking and closed-loop drone control [1]–[3]. Recent development of the Internet of Things (IoT) also promises real-time communication between numerous devices [4].

Since stale data is often of less value, it is crucial to maintain the *data freshness* of the system. An elementary approach is to transmit as many updates as possible. This, however, may clog the network [5] and consume excessive energy [6]. Recently, Age-of-Information (AoI) was introduced to characterize the level of information freshness [7], which has since been the foundation of many studies on data freshness control [8].

Early AoI minimization works [9]–[11] studied the model where update packets arrive at the destination according to specific stochastic processes. Various queueing models from $M/M/1$ to $M/G/1$ were considered and the closed-form expression of the average AoI was derived and minimized in [12], [13]. [14] proposed the *generate-at-will* model and showed that to minimize the average AoI, the source node often has to *wait before sending the next packet* even when the channel/queue is currently idle.

Generalizing [14], Sun *et al.* [15], [16] considered the 1-way delay model, where random delay exists in the source-to-destination direction while the destination-to-source ACK is instantaneous, and characterized the optimal waiting time. [17] later showed that the transmitter should employ a threshold-based waiting policy in a different but related context of remote estimation of a Wiener process.

This work revisits and significantly strengthens the existing results [15], [16] with the following main contributions: (i) *Generalization from the 1-way delay to the* 2-way delay *setting,* i.e., the ACK also experiences delay. For comparison, almost all[1] existing works [12], [13], [15]–[17], [19]–[22] considered 1-way delay with instantaneous ACK. The generalization to 2-way delay will significantly broaden the applications of the theoretic AoI results to countless many practical scenarios in which the forward and backward delays are of comparable magnitude.

(ii) All existing results [15]–[18] used a bisection search to find the optimal policy, which is known to exhibit linear convergence. In contrast, *we propose a new way of computing the optimal policy with* quadratic convergence*, an order-of-magnitude improvement over the state of the art.*

(iii) In all prior works [12], [13], [15]–[22] except [23],[2] the knowledge of exact probability distribution of delay is required before one can numerically find the optimal waiting policy. On the surface, this requirement seems to be indispensable since the main goal of AoI minimization is to optimally adjust the waiting time to "match" the underlying delay distribution. Nonetheless, in practice it may be difficult if not impossible to know the underlying delay distribution[3] *a priori* since the delay distribution is constantly subject to network topology changes and traffic fluctuations [24], [25]. *This work derives a new low-complexity adaptive online algorithm that provably converges to the optimal policy without knowing the exact de-*

[1] [18] also considered 2-way delay. For comparison, the main focus of [18] was to unify AoI minimization [15], [16] and remote estimation [17]. A more traditional dynamic-programming-based solution was proposed in [18], which is fundamentally different from the fixed-point equation analysis in this work. Furthermore, [18] relied on the linearly-convergent bisection method and required the complete knowledge of the delay distribution. Both points are remarkably improved in this work.

[2] [23] proposed a reinforcement learning (RL) approach to learn the waiting time *without knowing the delay distribution*. While exhibiting some promising performance, RL is not able to converge to the optimal scheme in any of the simulation in [23], which consists of both the exponential and log-normal delays. In contrast, this work proposes a provably optimal adaptive scheme.

[3] Even the task of estimating the delay distribution can be time-consuming since each sample (transmission) takes a full round-trip-time to complete and one may need many samples to accurately estimate the probability density function.

This work was supported in parts by NSF under Grant CCF-1422997, Grant CCF-1618475, and Grant CCF-1816013.

lay distribution, a surprising result that could have substantial impact to practical AoI minimization protocol design.

II. Model and Formulation

A. System Model with Two-way Delay

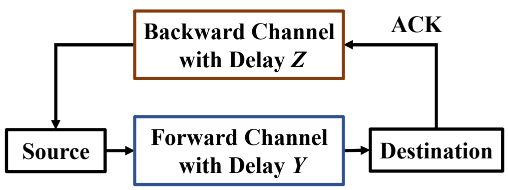

Fig. 1: Our system model with two-way delay.

Consider the system in Fig. 1, which comprises a source, a destination, a forward source-to-destination (s2d) channel and a backward destination-to-source (d2s) channel. We assume the following ACK-based model: At any time instant $t \in \mathbb{R}_+$, the source can generate a (status) update packet and transmit it to the destination, the *generate-at-will* model [14]–[16]. After the transmission, the source enters a listening mode and waits for the ACK from the destination. Once the source receives the ACK, then it can either transmit the next update packet immediately or wait for an arbitrary amount of time. After the next transmission, it again enters the listening mode and waits for ACK. The process repeats itself indefinitely.

Whenever the destination receives an update packet, an ACK is transmitted back to the source immediately. Both the s2d and d2s channels incur some random delay. We also assume all packets are time stamped. We now describe the detailed system evolution as follows.

Time sequences: The system consists of three discrete-time real-valued non-negative random processes X_i, Y_i, and Z_i, for all $i \geq 0$. X_i is the waiting time of the i-th update packet at the source;[4] Y_i (resp. Z_i) is the random delay for the i-th use of the s2d (resp. d2s) channel.

The instant when the i-th waiting time is over is denoted by S_i. That is, at time S_i, the i-th packet is generated and immediately transmitted. It is delivered to the receiver at time D_i. The source will receive its ACK at time A_i. The values of (S_i, D_i, A_i) refer to the absolute time instants while the values of (X_i, Y_i, Z_i) represent the lengths of the intervals. They are related by the following equations: Initialize $A_0 = X_0 = Y_0 = Z_0 = 0$. For all $i \geq 1$, we have $S_i = A_{i-1} + X_i$, $D_i = S_i + Y_i$, and $A_i = D_i + Z_i$. We call the time interval $[A_{i-1}, A_i)$ as the *i-th round*, which consists of the i-th waiting time X_i at the source, the i-th forward delay Y_i and backward delay Z_i. See Fig. 2 for illustration.

Age-of-Information and its penalty: Following [7], we define the Age-of-Information $\Delta(t)$ at time t by

$$\Delta(t) \triangleq t - \max\{S_i : i \text{ satisfies } D_i \leq t\}. \quad (1)$$

[4]As in most TCP-based control protocols [26], this setting prohibits the source from transmission before receiving the ACK (i.e., $X_i \geq 0$). One may design an even better scheme that transmits anticipatively before ACK is received, which, however, is beyond the scope of this work.

Let $\gamma(\cdot) : [0, \infty) \to [0, \infty)$ be a continuous, non-negative, and non-decreasing penalty function satisfying $\gamma(0) = 0$. We use $\gamma(\Delta(t))$ to represent the level of data staleness. Three popular choices are: (i) linear $\gamma_{\text{lin}}(\Delta) = \Delta$ [27]; (ii) exponential $\gamma_{\text{exp}}(\Delta) = e^{a\Delta} - 1$ for some constant $a > 0$ [16]; and (iii) quadratic $\gamma_{\text{qdr}}(\Delta) = \Delta^2$ [15]. Our results hold for any choice of $\gamma(\cdot)$. The evolution of $\gamma(\Delta(t))$ is plotted in Fig. 2.

Technical assumptions: We assume (i) Y and Z are of *bounded support*; (ii) (Y_i, Z_i) can be of arbitrary joint distribution $\mathbb{P}_{YZ}$ but the vector random process $\{(Y_i, Z_i) : i \geq 1\}$ is stationary and memoryless; and (iii) $\mathbb{E}\{Y_i\} + \mathbb{E}\{Z_i\} > 0$.

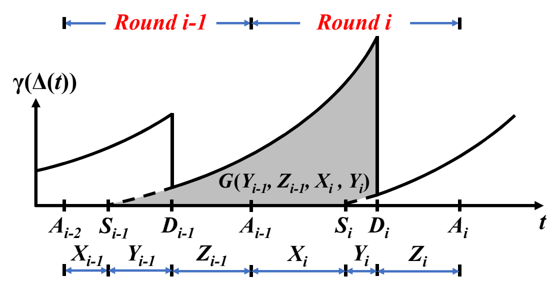

Fig. 2: Evolution of the AoI penalty function $\gamma(\Delta(t))$.

B. The Objective

Our goal is to minimize the long-term average AoI penalty:

$$\beta^* \triangleq \min_{\{X_i\}} \limsup_{T \to \infty} \frac{1}{T} \int_0^T \mathbb{E}\{\gamma(\Delta(t))\}\, dt. \quad (2)$$

We now define two deterministic functions:

$$G(y', z', x, y) \triangleq \int_0^{y'+z'+x+y} \gamma(t)dt - \int_0^{y} \gamma(t)dt \quad (3)$$

$$G_1(y', z', x) \triangleq \mathbb{E}_Y\{G(y', z', x, Y)\} \quad (4)$$

where $G_1(y', z', x)$ is the expectation of $G(y', z', x, Y)$ over Y. The intuition behind defining (3) is that the shaded area in Fig. 2 is characterized by $G(Y_{i-1}, Z_{i-1}, X_i, Y_i)$. By noticing that the overall area underneath $\gamma(\Delta(t))$ can be decomposed as a summation of smaller sub-areas with shapes similar to the shaded area $G(Y_{i-1}, Z_{i-1}, X_i, Y_i)$ in Fig. 2, we have

$$\beta^* = \min_{\{X_i\}} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} \mathbb{E}\{G(Y_{i-1}, Z_{i-1}, X_i, Y_i)\}}{\sum_{i=1}^{n} \mathbb{E}\{Y_{i-1} + Z_{i-1} + X_i\}}. \quad (5)$$

Since (5) is a Markov decision problem with i.i.d. $\{(Y_i, Z_i)\}$, it suffices to find the optimal policy for the *single-round* optimization problem instead (see [15], [16] for the detailed derivation). The optimization problem (5) can thus be simplified as

$$\beta^* = \min_{X_i} \frac{\mathbb{E}\{G_1(Y_{i-1}, Z_{i-1}, X_i)\}}{\mathbb{E}\{Y_{i-1} + Z_{i-1} + X_i\}} \quad (6)$$

where the numerator of (6) follows from (4) and from Y_i being independent of (Y_{i-1}, Z_{i-1}, X_i).

III. Main Results

A. Optimal Hitting-time-based Policy

At time A_{i-1}, the source has the knowledge of the past delays Y_{i-1} and Z_{i-1} since all packets are time stamped. As a result, we can write any waiting time rule $X_i = \phi(Y_{i-1}, Z_{i-1})$ as a function of (Y_{i-1}, Z_{i-1}). The resulting[5] averaged AoI penalty, not necessarily the minimum one, becomes

$$\text{Avg. AoI Penalty: } \frac{\mathbb{E}\left\{G_1\left(Y_{i-1}, Z_{i-1}, \phi(Y_{i-1}, Z_{i-1})\right)\right\}}{\mathbb{E}\left\{Y_{i-1} + Z_{i-1} + \phi(Y_{i-1}, Z_{i-1})\right\}}. \quad (7)$$

We now describe a special scheme. For any given $\beta > 0$, the scheme Γ_β has the following special decision rule:

$$X_i = \phi_{\Gamma,\beta}(Y_{i-1}, Z_{i-1}) \quad (8)$$

$$\triangleq \inf\left\{t > 0 : \frac{d}{dt} G_1(Y_{i-1}, Z_{i-1}, t) > \beta\right\}. \quad (9)$$

By (4), $G_1(Y_{i-1}, Z_{i-1}, t)$ is the conditional expectation (given (Y_{i-1}, Z_{i-1})) of the expected AoI penalty (the shaded area in Fig. 2) if the i-th waiting time is $X_i = t$. Therefore, the decision rule $\phi_{\Gamma,\beta}$ essentially chooses the hitting time for which the *growth rate*[6] of the conditional expected AoI penalty $G_1(Y_{i-1}, Z_{i-1}, t)$ first hits the threshold β.

For this scheme Γ_β, we use $f_\Gamma(\beta)$ to denote its average AoI penalty, which can be computed by substituting the ϕ in (7) with the $\phi_{\Gamma,\beta}$ in (8). The input argument "(β)" highlights the fact that the average AoI penalty of the decision rule $\phi_{\Gamma,\beta}$ is a function of the hitting time threshold β. Before proceeding, we introduce a simple lemma without its proof.

Lemma 1: For any positive constants p_1, T_1, r_1, p_2, T_2, r_2, τ, $r_\tau > 0$, we have the following two "$\Longrightarrow$" statements:

$$\frac{p_1 T_1 r_1 + p_2(T_2 r_2 + \tau r_\tau)}{p_1 T_1 + p_2(T_2 + \tau)} \le r_\tau \quad (10)$$

$$\Longrightarrow \frac{p_1 T_1 r_1 + p_2 T_2 r_2}{p_1 T_1 + p_2 T_2} \le \frac{p_1 T_1 r_1 + p_2(T_2 r_2 + \tau r_\tau)}{p_1 T_1 + p_2(T_2 + \tau)} \quad (11)$$

and

$$\frac{p_1 T_1 r_1 + p_2 T_2 r_2}{p_1 T_1 + p_2 T_2} \ge r_\tau \quad (12)$$

$$\Longrightarrow \frac{p_1 T_1 r_1 + p_2(T_2 r_2 + \tau r_\tau)}{p_1 T_1 + p_2(T_2 + \tau)} \le \frac{p_1 T_1 r_1 + p_2 T_2 r_2}{p_1 T_1 + p_2 T_2}. \quad (13)$$

Proposition 1: For any arbitrary scheme A with scheduling rule ϕ_A, we use β_A to denote its average AoI penalty, computed by substituting the ϕ in (7) with ϕ_A. The following inequality must hold: $f_\Gamma(\beta_A) \le \beta_A$.

That is, for any scheme A with average AoI penalty β_A, if we use β_A as the hitting time threshold in (9), then $f_\Gamma(\beta_A)$, the AoI penalty of the new scheme Γ_{β_A}, will be no worse than the average AoI penalty β_A of the original scheme A.

[5]The scheduling rule ϕ can be deterministic or randomized. In case of the latter, the expectation in (7) takes the average over the randomness in ϕ.

[6]If $G_1(Y_{i-1}, Z_{i-1}, t)$ is not differentiable, we can use the subgradient instead [28]. For simplicity, we assume differentiability.

Proof: Due to the space limit, we provide a high-level sketch of the proof. For schemes A and Γ_{β_A}, define S_i^A and S_i^Γ as the respective times when the i-th packet is transmitted. Suppose we are in the event of $S_i^\Gamma < S_i^A$, i.e., the scheme Γ_{β_A} sends the i-th update earlier than the scheme A. *During the interval* $\left[S_i^\Gamma, S_i^A\right]$, *the growth rate of* $G_1(Y_{i-1}, Z_{i-1}, t)$ *is higher than* β_A since S_i^Γ, as implied in (9), is the first time the growth rate hits β_A and the growth rate is non-decreasing (due to non-decreasing $\gamma(\cdot)$). Compared to the original scheme A, the new scheme Γ_{β_A} avoids "higher-than-β_A" average during the interval $\left[S_i^\Gamma, S_i^A\right]$, which in turn helps make its average AoI penalty $f_\Gamma(\beta_A)$ smaller than the benchmark β_A.

Mathematically speaking, average AoI penalty is the ratio of two expectations, see Lemma 1. In the left-hand side of (10), there is a duration of length τ that has the penalty growth rate r_τ larger than the current average, the inequality in (10). By avoiding this duration, the new average becomes the left-hand side of (11), which is better than the original average AoI, i.e., the inequality in (11).

Similarly, in the event of $S_i^A \le S_i^\Gamma$, during the interval $\left[S_i^A, S_i^\Gamma\right]$, scheme Γ_{β_A} will experience "lower-than-β_A" growth rate since the growth rate of $G_1(Y_{i-1}, Z_{i-1}, t)$ has not hit β_A yet, which again helps make $f_\Gamma(\beta_A)$ lower than β_A (as proved in (12) and (13)). Since in either case the average AoI penalty of Γ_{β_A} has improved over the benchmark β_A, we have Proposition 1. ■

Recall that β^* is the minimum of (6). Since Γ_{β^*} is yet another scheme, (6) implies $\beta^* \le f_\Gamma(\beta^*)$. On the other hand, Proposition 1 implies $\beta^* \ge f_\Gamma(\beta^*)$. Jointly we have

Corollary 1: β^* is a root of the fixed-point equation

$$\beta = f_\Gamma(\beta). \quad (14)$$

Furthermore, if we know the value of β^*, then we can obtain the optimal policy for the 2-way delay setting by plugging β^* into the hitting time rule $\phi_{\Gamma,\beta}(\cdot,\cdot)$ in (8) and (9).

Remark 1: Corollary 1 is similar to [15, Theorem 3] and [16, Theorem 1]. This is as expected since our 2-way setting collapses to those of [15], [16] when hardwiring $Z_i = 0$.

Remark 2: The way we derive Corollary 1 is new. In [15], [16], the authors first defined the corresponding Lagrangian, then reformulated and solved it as a convex optimization problem, and finally showed that it admits no duality gap. The analytical tools used include the extension of Dinkelbach's method and the geometric multiplier technique.[7] In contrast, we first prove an intuitive result in Proposition 1 and the optimality conditions then follow suit naturally.

Remark 3: The function $f_\Gamma(\beta)$ can be computed easily by (3), (4), (7), (8), (9), together with the complete knowledge of distribution $\mathbb{P}_{Y_{i-1}, Z_{i-1}}$.

B. Fast Fixed-point Iteration for Computing β^*

Lemma 2: The root of $\beta = f_\Gamma(\beta)$ is unique, regardless of how we choose the penalty function $\gamma(\cdot)$.

[7]The setting in [15], [16] includes the maximum update frequency constraint (MUFC), which is not considered in this work. It is possible that the MUFC mandates the use of more advanced analytical tools.

We omit the proof due to space limits. We now present a new way of computing β^* using (14).

Proposition 2: Assume a non-restrictive condition that $f_\Gamma(\beta)$ is doubly continuously differentiable.[8] Set $\beta_0 = 0$ and iteratively compute $\beta_i = f_\Gamma(\beta_{i-1})$ for all $i = 1, 2, 3, \cdots$. The resulting sequence $\{\beta_i : i \geq 1\}$ is non-increasing and converges to the optimal β^* with quadratic convergence speed.

Proof: For all i strictly larger than 0, we have

$$\beta_{i+1} = f_\Gamma(\beta_i) \leq \beta_i \tag{15}$$

where "$\leq$" follows from Proposition 1. $\{\beta_i : i \geq 1\}$ is thus non-increasing. Since $\beta_i \geq \beta^*$ for all $i \geq 1$, the sequence converges. Since $\lim_{i\to\infty} \beta_i$ must be a root of $\beta = f_\Gamma(\beta)$, Lemma 2 implies $\lim_{i\to\infty} \beta_i = \beta^*$. The quadratic convergence is established by proving $\forall i \geq 1$,

$$|\beta_{i+1} - \beta^*| \leq \left(\max_{z\in[\beta^*,\beta_1]} \frac{|f''_\Gamma(z)|}{2} \right) |\beta_i - \beta^*|^2. \tag{16}$$

To that end, we apply Taylor's theorem [29] to $f_\Gamma(\beta)$:

$$\begin{aligned} &\beta_{i+1} - \beta^* = f_\Gamma(\beta_i) - \beta^* \\ &= \left(f_\Gamma(\beta^*) + (\beta_i - \beta^*) f'_\Gamma(\beta^*) + \frac{f''_\Gamma(z)}{2} (\beta_i - \beta^*)^2 \right) - \beta^* \end{aligned}$$

for some $z \in [\beta^*, \beta_i]$. Then, by (i) $f_\Gamma(\beta^*) = \beta^*$ and (ii) $f'_\Gamma(\beta^*) = 0$ (since β^* minimizes $f_\Gamma(\beta)$, see Sec. III-A), we have

$$\beta_{i+1} - \beta^* = \frac{f''_\Gamma(z)}{2} (\beta_i - \beta^*)^2. \tag{17}$$

Eq. (17) implies that (16) holds for all $i \geq 1$. ■

IV. Distribution-oblivious Online Algorithm

The design of the distribution-oblivious online algorithm is much more involved and we thus omit all the proofs due to the limited space. Before proceeding, we first define

$$g_1(y', z', \beta) \triangleq G_1(y', z', \phi_{\Gamma,\beta}(y', z')) \tag{18}$$

$$g_2(y', z', \beta) \triangleq y' + z' + \phi_{\Gamma,\beta}(y', z') \tag{19}$$

$$\overline{g}_1(\beta) \triangleq \mathbb{E}_{Y_{i-1},Z_{i-1}} \{g_1(Y_{i-1}, Z_{i-1}, \beta)\} \tag{20}$$

$$\overline{g}_2(\beta) \triangleq \mathbb{E}_{Y_{i-1},Z_{i-1}} \{g_2(Y_{i-1}, Z_{i-1}, \beta)\} \tag{21}$$

Comparing these four definitions to (7) and recalling that $f_\Gamma(\beta)$ is defined as the average AoI penalty when $X_i = \phi_{\Gamma,\beta}(Y_{i-1}, Z_{i-1})$, it is clear that

$$f_\Gamma(\beta) = \frac{\overline{g}_1(\beta)}{\overline{g}_2(\beta)}. \tag{22}$$

[8]For instance, if Y_i and Z_i have finite support and γ is doubly continuously differentiable, then $f_\Gamma(\beta)$ is also doubly continuously differentiable.

A. Description of the Proposed Scheme

For any $i \geq 1$, at time A_{i-1}, the source learns the values of (Y_{i-1}, Z_{i-1}) from the time stamps in the received ACK, and computes a β_i value, in a way to be explained shortly. After β_i is computed, the source substitutes the β in (9) with β_i and computes the i-th waiting time $X_i(\beta_i)$, for which we use "(β_i)" to emphasize that β_i is used as the threshold.

At time $S_i = A_{i-1} + X_i(\beta_i)$, the source generates and transmits the update packet. The β_i used at time A_{i-1} will be iteratively computed according to the following formula.

First choose a sufficiently large[9] constant $\beta_{\max}$ that is guaranteed to be larger than β^*. Then set $\beta_1 = \beta_2 = 0$ and for all $i \geq 3$ set

$$\beta_i = \min \left(\frac{\sum_{j=1}^{i-1} g_1(Y_{j-1}, Z_{j-1}, \beta_j)}{\sum_{j=1}^{i-1} g_2(Y_{j-1}, Z_{j-1}, \beta_j)}, \beta_{\max} \right) \tag{23}$$

$$= \min \left(\frac{\sum_{j=1}^{i-1} g_1(Y_{j-1}, Z_{j-1}, \beta_j)}{S_{i-1}}, \beta_{\max} \right) \tag{24}$$

where each $g_1(Y_{j-1}, Z_{j-1}, \beta_j)$ in the numerator can be viewed as the empirical AoI penalty experienced during time interval (S_{j-1}, S_j). There is no need to repeat the summation $\sum_{j=1}^{i-1} g_1(Y_{j-1}, Z_{j-1}, \beta_j)$ for each i and we only need to "update" the sum by adding the increment from i to $i+1$. The denominator $\sum_{j=1}^{i-1} g_2(Y_{j-1}, Z_{j-1}, \beta_j) = S_{i-1}$ follows from

$$\begin{aligned} &g_2(Y_{j-1}, Z_{j-1}, \beta_j) = Y_{j-1} + Z_{j-1} + \phi_{\Gamma,\beta_j}(Y_{j-1}, Z_{j-1}) \\ &= Y_{j-1} + Z_{j-1} + X_j = S_j - S_{j-1}. \end{aligned} \tag{25}$$

Note that (24) computes the ratio of the past total AoI penalty over the past duration $[0, S_{i-1}]$, which is essentially the *empirical average AoI penalty*. We then use it as the new threshold β_i to decide the $X_i(\beta_i)$ for the i-th round. This closely follows the spirit of the fixed-point iteration

$$\beta_i = f_\Gamma(\beta_{i-1}) = \frac{\overline{g}_1(\beta_{i-1})}{\overline{g}_2(\beta_{i-1})} \tag{26}$$

in Proposition 2. The differences between (23) and (26) are (i) (23) not only depends on β_{i-1} but also on $\{\beta_j : j \leq i-1\}$ and (ii) (23) uses the empirical $g_1(Y_{j-1}, Z_{j-1}, \beta_j)$ and $g_2(Y_{j-1}, Z_{j-1}, \beta_j)$ rather than the expectations $\overline{g}_1(\beta_{i-1})$ and $\overline{g}_2(\beta_{i-1})$. Therefore, $\{\beta_i\}$ in (23) is a *random process* but $\{\beta_i\}$ in Proposition 2 is a deterministic sequence.

Proposition 3 (Convergence): There exists an $\alpha \in (0, 0.5)$ and four constants $c_1, c_2, c_3, c_4 > 0$ such that $\forall i \geq 1$,

$$\mathbb{P}\left(\beta_{i+1} < \beta^* - c_1 \cdot i^{-(0.5-\alpha)}\right) \leq c_2 \exp\left(-c_3 \cdot i^{2\alpha}\right) \tag{27}$$

$$\mathbb{E}\{\beta_i - \beta^*\} \leq c_4 \cdot i^{-(0.5-\alpha)} \tag{28}$$

Proposition 3 shows that the random process $\{\beta_i : i\}$ computed in (24) converges to β^* in probability.

[9]We introduce $\beta_{\max}$ for the rigor of the analysis, which prevents β_i from growing unboundedly. In practice, we can simply choose an extremely large $\beta_{\max}$, e.g., 10^9, which will have zero impact on how the algorithm runs.

B. Analyzing the Knowledge Required to Run the Algorithm

The source automatically knows the value of the denominator of (24) by its local clock S_{i-1}. Since the source knows the history of $(Y_{j-1}, Z_{j-1}, \beta_j)$, to compute the numerator of (24), the algorithm only needs to know how to compute the value of $g_1(y', z', \beta)$ for all (y', z', β) and then plugs in the historical values of $(Y_{j-1}, Z_{j-1}, \beta_j)$ along the way. However, a closer look at (3), (4), and (18) shows that

$$g_1(y', z', \beta) = \mathbb{E}_Y \left\{ \int_Y^{y'+z'+\phi_{\Gamma,\beta}(y',z')+Y} \gamma(t)dt \right\} \quad (29)$$

which still requires some knowledge of the distribution $\mathbb{P}_Y$. However, for the three popular choices of $\gamma(\Delta)$ in Sec. II-A, the computation of $g_1(y', z', \beta)$ takes little effort. For example, with linear $\gamma_{\text{lin}}(\Delta) = \Delta$, applying simple calculus to (3), (4), (9), and (18) shows that

$$\phi_{\Gamma,\beta}(y', z') = \max\left(\beta - \mathbb{E}\{Y\} - y' - z', 0\right) \quad (30)$$

$$\begin{aligned} g_1(y', z', \beta) =& \frac{(y' + z' + \phi_{\Gamma,\beta}(y', z'))^2}{2} \\ &+ (y' + z' + \phi_{\Gamma,\beta}(y', z'))\,\mathbb{E}\{Y\} \end{aligned} \quad (31)$$

which requires only the knowledge of $\mathbb{E}\{Y\}$. In practice, we can replace $\mathbb{E}\{Y\}$ with the running empirical average $\frac{\sum_{j=1}^{i-1} Y_j}{i-1}$ and we thus have a truly distribution-oblivious online algorithm. Similarly, with exponential $\gamma_{\text{exp}}(\Delta)$, one only needs $\mathbb{E}\{e^{aY}\}$ to compute $g_1(y', z', \beta)$. With quadratic $\gamma_{\text{qrd}}(\Delta)$, one only requires $\mathbb{E}\{Y\}$ and $\mathbb{E}\{Y^2\}$ to compute $g_1(y', z', \beta)$. Again, we can use running empirical average as a substitute when executing the algorithm.

V. Simulation Results

Fig. 3 presents our numerical results of (i) the distribution-oblivious online algorithm and (ii) fixed-point-iteration-based computation of β^*. Exponential and log-normal delays are used since they are empirically reasonable channel models [30], [31]. In Sec. IV-B, we discuss the difference of still requiring $\mathbb{E}\{Y\}$ versus using the running empirical average as a substitute. In our simulation results, there is no visible difference between the two versions and we thus report only the results using the running empirical average of $\mathbb{E}\{Y\}$, $\mathbb{E}\{Y^2\}$ and $\mathbb{E}\{e^{aY}\}$ when computing $g_1(y', z', \beta)$. In other words, the results of the truly distribution-oblivious algorithms.

First, we consider independent exponential delays with $\lambda_Y = \lambda_Z = 0.2$ and linear AoI penalty $\gamma_{\text{lin}}(\Delta) = \Delta$. Fig. 3a plots the evolution of β_i versus i and benchmarks β_i against β^* (the red dashed line). The three curves in Fig. 3a are generated by different random seeds. For each curve, β_i is within 6% of β^* after just 10^2 iterations. Since it is an online algorithm, *it means that using our distribution-oblivious scheme, after sending just 100 update packets, the average AoI penalty of the underlying system (over the last 100 packets) is already within 6% of the best offline solution that requires complete knowledge of the delay distributions.* The gap is less than 2% after 10^4 iterations. The behavior is consistent with the analytical convergence results in Proposition 3.

The trajectories of the fixed-point computation $\beta_{i+1} = f_\Gamma(\beta_i)$ versus the bisection method are plotted in Fig. 3b. The advantage of our scheme is twofold. Firstly it is faster than the bisection method.[10] Secondly, as proved in Proposition 2, the sequence $\{\beta_i\}$ is non-increasing and thus does not fluctuate as in the case of the bisection search.

Almost identical behaviors can be observed when changing the linear penalty to exponential penalty $\gamma_{\text{exp}}(\Delta) = e^{(\Delta/10)} - 1$ while using the same Y and Z distributions, see Figs. 3c and 3d. For quadratic AoI penalty $\gamma_{\text{qrd}}(\Delta) = \Delta^2$ with Y_i and Z_i being independent log-normal with $(\mu_Y, \sigma_Y^2) = (0.5, 0.25)$ and $(\mu_Z, \sigma_Z^2) = (0.5, 0.5)$ see Figs. 3e and 3f. Our distribution-oblivious algorithm always converges to the optimal value, and the fixed-point computation outperforms the bisection method, as expected by Proposition 2 and 3.

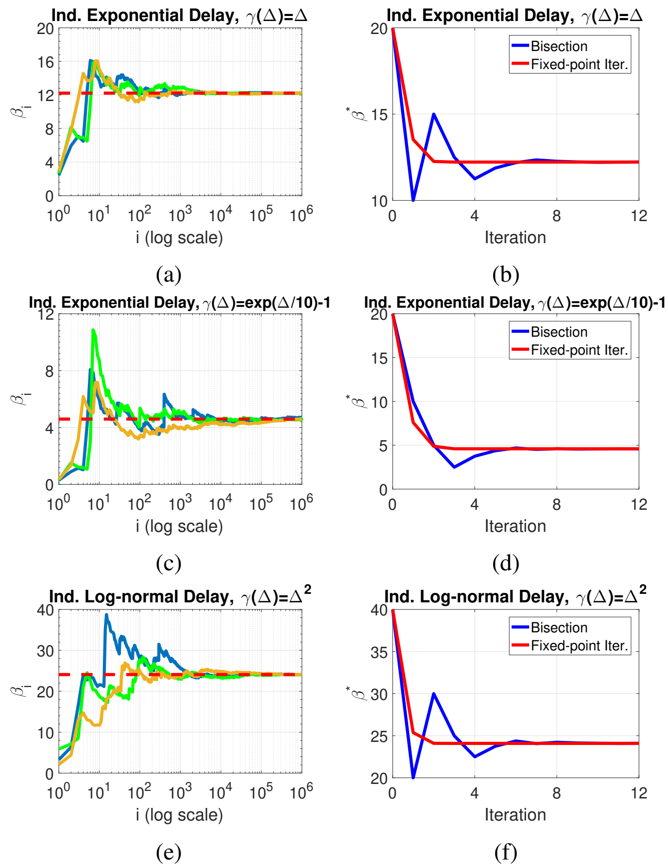

Fig. 3: Left: Evolution of β_i using the online algorithm (red dashed line for β^*). Right: The offline computation of β^*.

VI. Conclusion

We have proposed a new 2-way-delay AoI minimization framework, and derived the corresponding optimal waiting policy with quadratic convergence. We have also developed the first provably optimal distribution-oblivious online algorithm on AoI minimization.

[10] The closer β_i is to β^*, the greater the convergence speed improvement since one is quadratic and the other is linear.

REFERENCES

[1] S. Majumder, T. Mondal, and M. J. Deen, "Wearable sensors for remote health monitoring," *Sensors*, vol. 17, no. 1, p. 130, 2017.

[2] K. G.-A. Mataram and B. E. P.-S. N. Mandiri, "Implementation of Location Base Service on Tourism Places in West Nusa Tenggara by using Smartphone," *Publikasi Internasional*, vol. 1, no. 1, 2015.

[3] J. Fleureau, Q. Galvane, F.-L. Tariolle, and P. Guillotel, "Generic drone control platform for autonomous capture of cinema scenes," in *Proceedings of the 2nd Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*. ACM, 2016, pp. 35–40.

[4] A. V. Dastjerdi and R. Buyya, "Fog computing: Helping the Internet of Things realize its potential," *Computer*, vol. 49, no. 8, pp. 112–116, 2016.

[5] J. Korhonen, *Introduction to 3G mobile communications*. Artech House, 2003.

[6] L. Gavrilovska, S. Krco, V. Milutinović, I. Stojmenovic, and R. Trobec, *Application and multidisciplinary aspects of wireless sensor networks: concepts, integration, and case studies*. Springer Science & Business Media, 2010.

[7] X. Song and J. W.-S. Liu, "Performance of multiversion concurrency control algorithms in maintaining temporal consistency," in *Proceedings., Fourteenth Annual International Computer Software and Applications Conference*. IEEE, 1990, pp. 132–139.

[8] A. Kosta, N. Pappas, V. Angelakis *et al.*, "Age of information: A new concept, metric, and tool," *Foundations and Trends® in Networking*, vol. 12, no. 3, pp. 162–259, 2017.

[9] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *2012 46th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, 2012, pp. 1–6.

[10] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *2012 IEEE International Symposium on Information Theory Proceedings*. IEEE, 2012, pp. 2666–2670.

[11] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *2013 IEEE International Symposium on Information Theory*. IEEE, 2013, pp. 66–70.

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

[13] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1681–1685.

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

[15] 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.

[16] Y. Sun and B. Cyr, "Sampling for data freshness optimization: Non-linear age functions," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 204–219, 2019.

[17] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the Wiener process over a channel with random delay," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 321–325.

[18] C.-H. Tsai and C.-C. Wang, "Unifying AoI Minimization and Remote Estimation — Optimal Sensor/Controller Coordination with Random Two-way Delay," in *IEEE INFOCOM 2020-IEEE International Conference on Computer Communications*. IEEE, April 2020, accepted and to appear.

[19] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *2014 IEEE International Symposium on Information Theory*. IEEE, 2014, pp. 1583–1587.

[20] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *2015 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2015, pp. 1681–1685.

[21] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *2016 IEEE International Symposium on Information Theory (ISIT)*. Ieee, 2016, pp. 2574–2578.

[22] K. Chen and L. Huang, "Age-of-information in the presence of error," in *2016 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2016, pp. 2579–2583.

[23] C. Kam, S. Kompella, and A. Ephremides, "Learning to sample a signal through an unknown system for minimum AoI," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 177–182.

[24] H. Pucha, Y. Zhang, Z. M. Mao, and Y. C. Hu, "Understanding network delay changes caused by routing events," in *ACM SIGMETRICS performance evaluation review*, vol. 35, no. 1. ACM, 2007, pp. 73–84.

[25] F. Wang, Z. M. Mao, J. Wang, L. Gao, and R. Bush, "A measurement study on the impact of routing events on end-to-end internet path performance," in *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4. ACM, 2006, pp. 375–386.

[26] D. Lee, B. E. Carpenter, and N. Brownlee, "Media streaming observations: Trends in UDP to TCP ratio," *International Journal on Advances in Systems and Measurements*, vol. 3, no. 3-4, 2010.

[27] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 486–493.

[28] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[29] R. G. Bartle and D. R. Sherbert, *Introduction to real analysis*. Wiley New York, 1992, vol. 2.

[30] L. Kleinrock, *Queueing systems, volume 2: Computer applications*. Wiley-Interscience, 1976.

[31] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 494–502, 1998.