Ao^{2}I: Minimizing Age of Outdated Information to Improve Freshness in Data Collection

Qingyu Liu[*] Chengzhang Li[*] Y. Thomas Hou[*] Wenjing Lou[*] Jeffrey H. Reed[*] Sastry Kompella[†]

[*]Virginia Tech, Blacksburg, VA, USA

[†]U.S. Naval Research Laboratory, Washington, DC, USA

***Abstract*—Recently, it has been recognized that there is a serious limitation with the original Age of Information (AoI) metric in terms of quantifying true freshness of information content. A new metric, called Age of Incorrect Information (AoII), has been proposed. By further refining this new metric with practical considerations, we introduce Age of Outdated Information (Ao^{2}I) metric. In this paper, we investigate a scheduling problem for minimizing Ao^{2}I in an IoT data collection network. We derive a theoretical lower bound for the minimum Ao^{2}I that any scheduler can achieve. Then we present Heh—a low-complexity online scheduler. The design of Heh is based on the estimation of a novel offline scheduling priority metric in the absence of knowledge of the future. We prove that at each time, transmitting one source with the largest offline scheduling priority metric minimizes Ao^{2}I. Through extensive simulations, we show that the lower bound is very tight and that the Ao^{2}I obtained by Heh is close-to-optimal.**

***Index Terms*—Age of Information, Age of Outdated Information, Age of Incorrect Information, Scheduling**

I. Introduction

Age of Information (AoI) is an application layer metric that has been used to estimate freshness of information [1], [2]. It measures the elapsed time between the present and the time when the stored information was initially generated at its source. Since its inception, AoI has attracted active research efforts due to its unique characterization of latency from an application perspective (see [3] for a comprehensive bibliography). A main line of AoI research has been focused on minimizing AoI under certain resource constraints, with resource ranges from bandwidth (e.g., [4]–[8]), energy (e.g. [9]–[12]), to throughput (e.g., [13]–[16]). Another line of research has been focused on modeling, analysis, and optimization of AoI (e.g., [17]–[19]). There are also some other branches of AoI research, such as scheduling subject to AoI constraints (e.g., [20]–[22]), game theory for AoI (e.g., [23]–[25]), and AoI applications (e.g., [26], [27]), to name a few.

As the research community accumulates more knowledge on AoI, it starts to realize its limitation in quantifying freshness of information. The fundamental issue here resides in the definition of *age*. Should the measure of age starts from the generation time of the information or from the first time instance when the information content changes? For example, if the time for information collected from its source has been elapsed for, say one day, and that there has been no new update at the source during this period, does this information still considered fresh or outdated? The original definition of AoI only accounts for the lapsed time since the information was initially generated, rather than how long the information has been outdated. But for most real-world applications, information with a large AoI may still be considered "fresh" if there are no new updates of information at the source. In other word, the existing definition of AoI does *not* accurately capture the true *freshness* of information from *content* perspective.

There have been some attempts to improve the original AoI metric definition, such as Age of Synchronization (AoS) [28] and Effective AoI (EAoI) [29].[1] AoS was introduced by Zhong *et al.* for a distributed content caching system [28]. Unlike AoI (which increases linearly), AoS remains zero when the source has no samples to send; similar to AoI, AoS grows linearly with time when the source collects a sample. Like AoI, AoS focuses on the time of sample collection at the source, and does not take into account the information content in the sample. EAoI was introduced by Yin *et al.* for proactive information update (from users' perspective) where a single server serves multiple users and the server may transmit information to a user only after the user makes a request [29]. Unlike AoI (which increases linearly), EAoI remains zero when the user does not make a transmission request; similar to AoI, EAoI grows linearly with time once the user makes a query for transmitting information. Similar to AoS, EAoI does not consider the information content in the sample.

Recently, Maatouk *et al.* [31] introduced a new metric called *Age of Incorrect Information (AoII)*. AoII measures the elapsed time between the present and the most recent time when the content of the sample at its source is the same as the content of the sample stored at the destination.[2] By definition, AoII will increase linearly with time but will drop under two events: (i) when the destination receives a new sample, and (ii) when the destination receives an indication[3] that its stored sample contains the same information as the latest sample at the source. AoII represents a major leap forward in AoI metric as it focuses on quantifying information freshness through the actual content in the information.

[1]Another metric called Generalized AoI (GAoI)—introduced by Feng *et al.* [30], measures uncertainty of collected samples. Since GAoI does not measure freshness of information from latency perspective, we omit its discussion in this paper.

[2]Although a general definition of AoII (involving product of two functions) exists in [31], it has not been well understood. To date, only a special case of the general definition has been studied [31]–[34]. Therefore, by AoII in this paper, we refer to the special case that is well understood.

[3]Maatouk *et al.* [31] assume such an indication can be sent from a source to the BS with zero latency.

Following the new AoII definition in [31], Maatouk *et al.* [31] studied an AoII minimization problem in a transmitter-receiver pair scenario, with an optimal scheduling policy developed. Chen and Ephremides [32] proposed an optimal scheduling policy for the same AoII minimization problem in [31] but considering a different model for source's sampling process. In [33], Kam *et al.* presented extensive numerical results to compare metrics of real-time error, AoI, and AoII under different scheduling policies. Kriouile and Assaad [34] designed a heuristic scheduling policy for minimizing AoII in a scenario where multiple users send status updates to a central entity.

Although results from [31]–[34] offer some early understandings on AoII, they share some serious limitations. They all assume a small and known state space for information content. This is because by definition, the AoII metric is too sensitive to information content. Hence mathematically, AoII is only tractable in very simple settings where (i) the state space for information content is small and known; and (ii) once the content of information changes, there is a reasonable probability that it will return to the previous state in the near future. However, in reality the state space for information content is usually very large (possibly infinite and unknown) and the likelihood of information content returning to a previous state in the near future is very small.

Inspired by the idea of AoII in [31], this paper aims to address the limitation of the state-of-the-art through a number of new advancements. First, we improve the AoII metric by introducing *Age of Outdated Information (Ao^{2}I)*. Ao^{2}I measure the elapsed time between the present and the first time when the stored information at the destination becomes outdated (changed) when compared at its source. Comparing to AoII, Ao^{2}I is imminently suitable to work with a very large (usually unknown) state space for information content, which is one would likely encounter in the real world and is more general than that considered by [31]–[34]. Using this improved metric, we investigate a scheduling problem for minimizing Ao^{2}I in an IoT network. This is a canonical scenario considered widely in AoI research (such as [4], [5], [7], [8], [13], [16], [20]–[22], [35]) and has a wide range of applications for data collection. The main contributions of this paper are summarized as follows:

- We develop a theoretical lower bound for the minimum Ao^{2}I that can be used as a benchmark for any scheduler. To do this, we first develop a connection between a source's Ao^{2}I and its data rate. Leveraging this connection, we replace the Ao^{2}I objective in the original scheduling problem with a data rate objective. Subsequently, we obtain a new optimization problem for minimizing data rates which is convex. The optimal solution of the reformulated problem gives a lower bound for the minimum Ao^{2}I that any scheduler can achieve.
- To help develop an online scheduler, we first propose a novel priority metric for an offline optimal scheduler. We prove that the offline scheduler that transmits the source with the largest priority metric in each time slot is optimal for minimizing Ao^{2}I. Then we present Heh—a low-complexity online scheduler. The essence of Heh is to first estimate the scheduling priority metric based on past history without any knowledge of the future, and then schedule the source with the largest estimated priority metric for transmission in each time slot.
- Through extensive simulations, we find that Ao^{2}I obtained by Heh is very close to the lower bound. This indicates that (i) Heh is near-optimal for minimizing Ao^{2}I, and (ii) the lower bound is very tight (very close to the minimum Ao^{2}I).

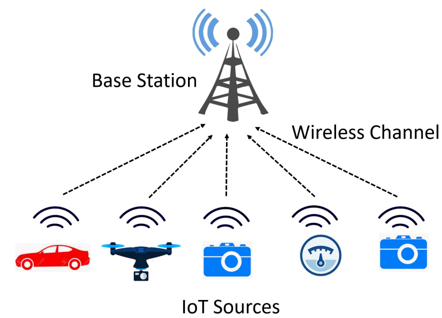


Fig. 1. Reference model: N sources collect samples from the environment and forward them to a BS through a shared wireless channel.

The rest of the paper is organized as follows. In Section II, we describe the system model that we will use in this paper to study Ao^{2}I. We also formally define Ao^{2}I mathematically and introduce the Ao^{2}I scheduling problem. In Section III, we derive a lower bound for the minimum Ao^{2}I. In Section IV, we present Heh, a low-complexity online scheduler that minimizes Ao^{2}I. In Section V, we present results from our simulation experiments. Section VI concludes this paper.

II. Problem Statement

Consider a scenario where there are N data collection sources and one receiving base station (BS) as shown in Fig. 1. Each source collects data samples from its environment and forwards them to the BS through a common (shared) wireless data channel. We assume (uplink) transmission time is slotted and each source collects a data sample at each time slot (see, e.g., [7], [8], [13], [16], [20]–[22], [35]). When a source is chosen for transmission at time t,[4] it will only transmit its freshest sample, i.e., the sample collected at time t, to the BS over the wireless data channel. At the BS side, it only maintains (stores) the sample that it has most recently received from each source. In this research, we assume: 1) the transmission of a sample takes exactly one time slot; and 2) at most one sample can be transmitted over the data channel in each time slot. This means that at most one source may transmit in each time slot.

A. Ao^{2}I Definition and Notations

Denote $A_i^B(t)$ as the AoI of source i at the BS at time t:

$$A_i^B(t) \;=\; t - G_i^B(t), \tag{1}$$

[4]For clarity, we use the term "at time t" to refer to "at the beginning of time slot t" and use the term "in time slot t" to refer to the underlying action is completed "at the end of time slot t".

where $G_i^B(t)$ is the generation time of the sample from source i that is currently (at time t) stored at the BS.

As discussed in Section I, AoI only accounts for raw lapsed time since the sample was initially generated, but not how long the *information content* in the sample has been outdated. In many scenarios, a sample with a large AoI may still contain the freshest information if there is no change of information content in the new samples at the source.

To address this issue, Maatouk *et al.* [31] proposed a new metric called AoII as follows. Denote $X_i^S(t)$ as the content of the sample collected by source i at time t. Denote $X_i^B(t)$ as the content of the sample that is from source i and stored at the BS at time t. The AoII of source i at the BS at time t is:

$$\Delta_i^B(t) \;=\; t - \max_{\hat{t}\le t,\ \hat{t}\in\mathbb{Z}^+} \left\{\hat{t} : X_i^S\left(\hat{t}\right) = X_i^B(t)\right\}, \qquad (2)$$

where $\mathbb{Z}^+$ is the set of positive integers. By definition, AoII measures the elapsed time between the present (t) and the most recent time when source i collects a sample that contains the same content as the sample stored at the BS at time t.

As discussed in Section I, the AoII $\Delta_i^B(t)$ is highly sensitive to $X_i^S(t)$ (i.e., $\Delta_i^B(t)$ drops at any time t if $X_i^S(t) = X_i^B(t)$). Hence existing AoII studies, including [31]–[34], indicate that mathematically $\Delta_i^B(t)$ is only tractable under following assumptions:

- The state space for $X_i^S(t)$'s is *small* and *known*;
- The transition probabilities between different states for $X_i^S(t)$'s is *known*;
- Once $X_i^S(t)$ changes (i.e., once $X_i^S(t) \neq X_i^S(t-1)$), there is a reasonable probability that it will return to the previous state in the near future.

But in reality, the state space for $X_i^S(t)$'s is usually very large (possibly infinite and unknown). As such, it is very difficult to obtain the transition probabilities among a large number of state space. Further, the likelihood of information content returning to a previous state in the near future is very small.

To improve AoII metric for more general scenarios and more amenable for mathematically analysis, we propose an improved metric called Ao^{2}I as follows. Denote $y_i^S(t_1, t_2)$ as an indicator of whether or not the sample content has changed in the time interval $[t_1, t_2]$ at source i, i.e.,

$$y_i^S(t_1,t_2) = \begin{cases} 1, & \text{if there exists a } \hat{t}\in[t_1,t_2], \hat{t}\in\mathbb{Z}^+ \\ & \quad \text{such that } X_i^S(\hat{t}) \neq X_i^S(t_1); \\ 0, & \text{otherwise}. \end{cases} \qquad (3)$$

Denote $U_i^S(t)$ as the largest time since $G_i^B(t)$ such that the information content of the sample remains *unchanged* at source i,[5] i.e.,

$$U_i^S(t) \;=\; \max_{G_i^B(t)\le\hat{t}<t,\ \hat{t}\in\mathbb{Z}^+} \left\{\hat{t} :\ y_i^S\left(G_i^B(t), \hat{t}\right) = 0\right\}. \qquad (4)$$

By definition in (4), i.e., $\hat{t} < t$ under the $\max$ function, we have $U_i^S(t) < t$. This is because we assume that it requires one time slot for the source to send a feedback to the BS to report

[5]Note that $U_i^S(t)$ is only defined for integer values of t in our discrete-time system model.

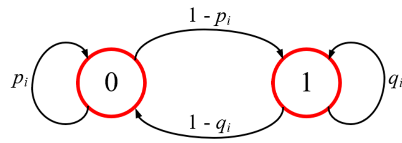


Fig. 2. A 2-state discrete-time Markov process to characterize the content change behavior at source i.

a change (details on the feedback procedure are presented later in Section II-C). Note that $U_i^S(t)$ is the last time when $X_i^B(t)$ is fresh at source i, and $\left(U_i^S(t)+1\right)$ is the first time when $X_i^B(t)$ is outdated at source i.[6] Denote $D_i^B(t)$ as the Ao^{2}I of source i at the BS at time t. Then we have:

$$D_i^B(t) \;=\; t - \left(U_i^S(t)+1\right). \qquad (5)$$

Ao^{2}I quantifies the elapsed time between the present (t) and the first time when the sample stored at the BS is changed (outdated) at its source ($U_i^S(t)+1$).

It is clear that the Ao^{2}I $D_i^B(t)$ addresses the limitation of AoII $\Delta_i^B(t)$. $\Delta_i^B(t)$ drops whenever the source collects a new sample that contains the same content as the sample currently stored at the BS, while $D_i^B(t)$ linearly increases with time after the first time when the sample currently stored at the BS is changed (outdated) at its source. In contrast to AoII, Ao^{2}I works best when the state space for information content is large (possibly unknown and infinite) and there is little probability that the information content will return to a previous state in the near future. Moreover, Ao^{2}I only requires the knowledge of transition probability between two states (change in information content or not) rather than transition probabilities among the states of information content, which are substantially harder to obtain. As we shall see shortly, Ao^{2}I is much more amenable to mathematical analysis and optimization than AoII.[7]

It is not hard to see that the Ao^{2}I $D_i^B(t)$ is fundamentally different from the AoI $A_i^B(t)$. $D_i^B(t)$ takes into account of the actual information content in the sample while $A_i^B(t)$ does not. As such, $D_i^B(t)$ will not increase with time as long as the sample content currently stored at the BS is not outdated. In contrast, $A_i^B(t)$ linearly increases with time when the BS is waiting to receive a new sample even though the sample content currently stored at the BS is not outdated. In this sense, Ao^{2}I is a much more accurate measure of freshness in information content than AoI.

B. Modeling Sample Content Update

In this paper, we assume the state (change or no change) of successive sample's content follows a 2-state discrete-time

[6]To be more accurate, (i) if $U_i^S(t) < t-1$, $U_i^S(t)+1$ is the first time when $X_i^B(t)$ is outdated; however, (ii) if $U_i^S(t) = t-1$, the BS does not know whether or not $U_i^S(t)+1 = t$ is the first time when $X_i^B(t)$ is outdated. This is because it requires one time slot for the source to send a feedback to the BS to report a change.

[7]Here we remark that both AoII and Ao^{2}I are only suitable when information content has *discrete* state space. In the case when the content follows a continuous state, e.g., a Brownian motion, one should first construct a discrete model (through discretization). Such discretization of continuous information content has been explored in [36].

Markov chain as depicted in Fig. 2. Denote $c_i^S(t) \in \{0,1\}$ as a binary state variable, indicating whether or not sample's content changes at source i at time t as compared to $t-1$:

$$c_i^S(t) = \begin{cases} 1, & \text{if } X_i^S(t) \neq X_i^S(t-1); \\ 0, & \text{otherwise}. \end{cases} \quad (6)$$

Denote p_i ($0 < p_i < 1$) and q_i ($0 < q_i < 1$) as state transition probabilities with the following definitions:

$$\mathbb{P}\left(c_i^S(t+1) = 0 | c_i^S(t) = 0\right) = p_i; \quad (7a)$$
$$\mathbb{P}\left(c_i^S(t+1) = 1 | c_i^S(t) = 0\right) = 1 - p_i; \quad (7b)$$
$$\mathbb{P}\left(c_i^S(t+1) = 1 | c_i^S(t) = 1\right) = q_i; \quad (7c)$$
$$\mathbb{P}\left(c_i^S(t+1) = 0 | c_i^S(t) = 1\right) = 1 - q_i. \quad (7d)$$

That is, if the content in the sample does not change from $t-1$ to t (i.e., $c_i^S(t) = 0$), then with a probability of p_i, the content will remain unchanged from t to $t+1$ (i.e. $c_i^S(t+1) = 0$) and so forth. Likewise, if the content in the sample changes from $t-1$ to t (i.e., $c_i^S(t) = 1$), then with a probability of q_i, the content will keep changing from t to $t+1$ (i.e., $c_i^S(t+1) = 1$) and so forth. It should be clear that the Markov model in Fig. 2 is independent of the state space of $X_i^S(t)$'s which can be very large (likely infinite) and unknown.

C. Feedback of Updates to the BS

An important question to ask is: How does the BS know any content change at source i (i.e., $c_i^S(t)$)? Clearly, obtaining such a knowledge requires a feedback from source i to the BS. Many feedback mechanisms may be employed for this purpose. Here we illustrate a simple 1-bit feedback approach.

Specifically, at each time t, source i knows the value of $c_i^S(t)$ after it collects a new sample. Then source i sends a 1-bit feedback $c_i^S(t)$ to the BS (over a dedicated control channel).[8] Given that only 1-bit is required in feedback, it can be received by the BS at time $t+1$.

D. Problem Statement and Technical Challenges

Denote $\bar{D}_i^B$ as the long-term average of source i's Ao^{2}I at the BS. We have:

$$\bar{D}_i^B = \mathbb{E}\left[\lim_{T\to\infty} \frac{1}{T} \cdot \sum_{t=1}^{T} D_i^B(t)\right]. \quad (8)$$

Denote $\bar{D}^B$ as the average Ao^{2}I over all N sources at the BS. We have :

$$\bar{D}^B = \frac{1}{N} \cdot \sum_{i=1}^{N} \bar{D}_i^B. \quad (9)$$

The BS needs a scheduler to decide which source should transmit in each time slot. Denote $\boldsymbol{S}(t)$—an $N \times 1$ vector—as the scheduling decision for time slot t, where the i-th element in $\boldsymbol{S}(t)$ is $S_i(t) \in \{0,1\}$. $S_i(t) = 1$ represents that source i will transmit a sample in t, and $S_i(t) = 0$ otherwise (i.e., no transmission for source i). We have

$$\sum_{i=1}^{N} S_i(t) \leq 1, \text{ for all } t \in \mathbb{Z}^+. \quad (10)$$

Our goal is to find an optimal scheduler $\boldsymbol{S}_{\text{OPT}}(t)$ that minimizes $\bar{D}^B$.[9] We denote $\bar{D}_{\text{OPT}}^B$ as the minimum (optimal) $\bar{D}^B$ that is achieved by $\boldsymbol{S}_{\text{OPT}}(t)$.

III. Performance Bound

In this section, we develop a theoretical lower bound for $\bar{D}_{\text{OPT}}^B$. This lower bound can be used as a benchmark to measure the performance of any scheduler for minimizing $\bar{D}^B$.

For a given scheduler $\boldsymbol{S}(t)$ with its i-th element $S_i(t) \in \{0,1\}$, $i = 1, 2, \cdots, N$, define $\bar{S}_i$ as:

$$\bar{S}_i = \lim_{T\to\infty} \frac{1}{T} \cdot \sum_{t=1}^{T} S_i(t). \quad (11)$$

Intuitively, $\bar{S}_i$ represents the long-term proportion of time that source i is scheduled for transmission (data rate of source i). Denote $\bar{\boldsymbol{S}}$ as the vector of $\bar{S}_i$'s. Similar to $\bar{\boldsymbol{S}}$, define $\bar{\boldsymbol{S}}_{\text{OPT}}$ for the optimal scheduler $\boldsymbol{S}_{\text{OPT}}(t)$, with its i-th element to be $\bar{S}_{\text{OPT},i}$.

In this section, we first develop a lower bound for $\bar{D}_{\text{OPT}}^B$ with the assumption that $\bar{\boldsymbol{S}}_{\text{OPT}}$ *was* known *a priori.* Then in Section III-B, we present a lower bound for $\bar{D}_{\text{OPT}}^B$ by removing this assumption.

A. A Lower Bound for $\bar{D}_{OPT}^B$ with Known $\bar{\boldsymbol{S}}_{OPT}$

Denote $\bar{D}_{\text{LB}}^B$ as a lower bound for $\bar{D}_{\text{OPT}}^B$. Given $\bar{\boldsymbol{S}}_{\text{OPT}}$, can we find $\bar{D}_{\text{LB}}^B$? The following lemma addresses this question.

Lemma 1: Given $\bar{\boldsymbol{S}}_{\text{OPT}}$, with $\bar{S}_{\text{OPT},i}$ ($i = 1, 2, \cdots, N$) being its i-th element, a lower bound for $\bar{D}_{\text{OPT}}^B$ is

$$\bar{D}_{\text{LB}}^B = \frac{1}{N} \sum_{i=1}^{N} \max\left\{0, \frac{1}{2}\left(\frac{1}{\bar{S}_{\text{OPT},i}} - (\bar{S}_{\text{OPT},i} + 2) \cdot r_i + 1\right)\right\}, \quad (12)$$

where

$$r_i = \max\left\{q_i + \frac{1-q_i}{p_i} \cdot \left(\frac{1}{1-p_i} + p_i - 1\right), \frac{1}{1-p_i}\right\}. \quad (13)$$

We offer a proof sketch here. A complete proof is given in Appendix A. Recall that $\bar{S}_{\text{OPT},i}$ is the proportion of time that source i is scheduled for transmission under the optimal solution $\boldsymbol{S}_{\text{OPT}}(t)$. Hence $\frac{1}{\bar{S}_{\text{OPT},i}}$ estimates the number of time slots between two consecutive transmissions from source i. By the Ao^{2}I definition, $D_i^B(t)$ should first remain 0 and then increase linearly with time for these $\frac{1}{\bar{S}_{\text{OPT},i}}$ time slots. Here r_i is the expected number of time slots for $D_i^B(t)$ to remain 0, which can be derived from the Markov model in Fig. 2. Therefore, by the end of these $\frac{1}{\bar{S}_{\text{OPT},i}}$ time slots, $D_i^B(t)$ is at least $(\frac{1}{\bar{S}_{\text{OPT},i}} - r_i)$. To get a lower bound for the average value of $D_i^B(t)$, for these $\frac{1}{\bar{S}_{\text{OPT},i}}$ time slots, we consider an extreme case where $D_i^B(t)$ remains 0 in the first r_i time slots, and

[8]To ensure reliability of feedback in a lossy channel, a reliable protocol may be employed (see, e.g., [37]).

[9]All of our results to this objective can be directly extended to another objective of minimizing the weighted sum $\sum_{i=1}^{N} w_i \cdot \bar{D}_i^B$.

increases from 1 to $(\frac{1}{\bar{S}_{\text{OPT},i}} - r_i)$ in the remaining $(\frac{1}{\bar{S}_{\text{OPT},i}} - r_i)$ time slots. In $\bar{D}^B_{\text{LB}}$ in (12), $\frac{1}{2}\left(\frac{1}{\bar{S}_{\text{OPT},i}} - (\bar{S}_{\text{OPT},i}+2)\cdot r_i + 1\right)$ is an estimate on $(1+2+\cdots+(\frac{1}{\bar{S}_{\text{OPT},i}} - r_i))/\frac{1}{\bar{S}_{\text{OPT},i}}$, i.e., an estimate on the $D_i^B(t)$ averaged over these $\frac{1}{\bar{S}_{\text{OPT},i}}$ time slots in the extreme case.

B. From Known $\bar{\boldsymbol{S}}_{OPT}$ to Unknown $\bar{\boldsymbol{S}}_{OPT}$

The $\bar{D}^B_{\text{LB}}$ given in Lemma 1 requires the knowledge of $\bar{\boldsymbol{S}}_{\text{OPT}}$. However, such knowledge is unavailable in practice. In this section, we aim to remove this assumption.

It is clear that since $\bar{D}^B_{\text{LB}}$ is a lower bound of $\bar{D}^B_{\text{OPT}}$, then a lower bound for $\bar{D}^B_{\text{LB}}$ must also be a lower bound for $\bar{D}^B_{\text{OPT}}$. Now we will try to find a lower bound for $\bar{D}^B_{\text{LB}}$ in (12) without requiring the knowledge of $\bar{\boldsymbol{S}}_{\text{OPT}}$.

We propose to solve a problem that minimizes the RHS of (12) subject to the constraints that $\bar{\boldsymbol{S}}_{\text{OPT}}$ is feasible. Then the optimal solution of this optimization problem will give us a lower bound for $\bar{D}^B_{\text{LB}}$ in (12). More formally, we consider the following problem OPT-LB:

$$\text{OPT-LB:} \qquad \min \sum_{i=1}^{N} \max\left\{0, \frac{1}{2}\left(\frac{1}{x_i} - (x_i+2)\cdot r_i + 1\right)\right\}$$

$$\text{s.t.} \ \sum_{i=1}^{N} x_i \le 1, \tag{14}$$

$$x_i \ge 0 \text{ for all } i = 1, 2, \cdots, N. \tag{15}$$

We argue that constraints (14) and (15) are necessary conditions for a scheduler (including $\boldsymbol{S}_{\text{OPT}}(t)$) to be feasible. This is because for any feasible scheduler $\boldsymbol{S}(t)$, from (11), we can get $\bar{\boldsymbol{S}}$ (a vector of $\bar{S}_i$'s). Clearly (15) holds for $\bar{\boldsymbol{S}}$. Further,

$$\begin{aligned}\sum_{i=1}^{N} \bar{S}_i &= \lim_{T\to\infty} \frac{1}{T}\cdot \sum_{i=1}^{N}\sum_{t=1}^{T} S_i(t) \\ &= \lim_{T\to\infty} \frac{1}{T}\cdot \sum_{t=1}^{T}\sum_{i=1}^{N} S_i(t) \le \lim_{T\to\infty} \frac{1}{T}\cdot \sum_{t=1}^{T} 1 = 1,\end{aligned}$$

indicating that (14) holds for $\bar{\boldsymbol{S}}$ since $\boldsymbol{S}(t)$ is a feasible scheduler and must satisfy the constraint (10). Therefore, by solving OPT-LB and finding an optimal solution to the x_i's (denoted as a vector $\boldsymbol{x}^*$ with x_i^* being its i-th element), we can find a lower bound for $\bar{D}^B_{\text{OPT}}$. The result is stated in the following lemma.

Lemma 2: A lower bound for $\bar{D}^B_{\text{OPT}}$ is

$$\bar{D}^B_{\text{LB}} = \frac{1}{N}\cdot \sum_{i=1}^{N} \max\left\{0, \frac{1}{2}\cdot\left(\frac{1}{x_i^*} - (x_i^*+2)\cdot r_i + 1\right)\right\}, \tag{16}$$

where r_i of each $i = 1, 2, \cdots, N$ is defined in (13) and $\boldsymbol{x}^*$ is the optimal solution to OPT-LB.

OPT-LB is a convex optimization problem and a commercial off-the-shelf (COTS) solver such as CVX [38] can be used to solve it efficiently. Clearly, OPT-LB does not require the knowledge of $\bar{\boldsymbol{S}}_{\text{OPT}}$.

IV. Heh: An Online Scheduler for Minimizing Ao^{2}I

In this section we present Heh[10]—an online scheduling algorithm for minimizing $\bar{D}^B$. At each time t, Heh decides which source to transmit based on p_i, q_i, $D_i^B(t)$, and $c_i^S(t-1)$ for $i = 1, 2, \cdots, N$. Note that at time t, the BS knows $c_i^S(t-1)$ but not $c_i^S(t)$ (see Section II-C).

A. Motivating Idea

At time t, if scheduling of source i's transmission can minimize $\bar{D}^B$ (as compared to scheduling of any other source's transmission), then source i should be selected for transmission. So the key question becomes: What criterion should we use to identify the source whose transmission will minimize $\bar{D}^B$ at time t? In the rest of this section, we develop such a criterion.

Recall that $\bar{D}^B$ is the average Ao^{2}I achieved by a specific scheduler, say $\boldsymbol{S}(\tau)$ $(\tau = 1, 2, \cdots)$. Also recall $S_i(t) \in \{0, 1\}$ is the i-th element of $\boldsymbol{S}(\tau)$ at time $\tau = t$. Now define $\bar{D}^B_{\boldsymbol{S}_{(t)}(\tau)}$ as the "modified" $\bar{D}^B$ that is achieved by the following scheduler $\boldsymbol{S}_{(t)}(\tau)$:

$$\boldsymbol{S}_{(t)}(\tau) = \begin{cases} \boldsymbol{S}(\tau) & \text{if } \tau < t; \\ \boldsymbol{R}(t), & \text{if } \tau = t; \\ \boldsymbol{S}(\tau), & \text{if } \tau > t. \end{cases} \tag{17}$$

That is, $\boldsymbol{S}_{(t)}(\tau)$ has the same scheduling behavior as $\boldsymbol{S}(\tau)$ except at time t when it follows scheduling $\boldsymbol{R}(t)$, which may differ from $\boldsymbol{S}(t)$. For ease of notation, we abbreviate $\bar{D}^B_{\boldsymbol{S}_{(t)}(\tau)}$ as $\bar{D}^B(\boldsymbol{R}(t))$ when there is no confusion.

With the above notation, we consider two different sources i and j that may be selected for transmission at time t.

- If at time t source i transmits, we have:

$$\begin{aligned}&\bar{D}^B(\boldsymbol{R}(t)\text{: } R_i(t) = 1, R_k(t) = 0 \text{ for } k \neq i) = \frac{1}{N}\cdot \\ &\left(\bar{D}_i^B(R_i(t) = 1) + \sum_{k=1,k\neq i}^{N} \bar{D}_k^B(R_k(t) = 0)\right).\end{aligned} \tag{18}$$

- Likewise, if at time t source j transmits, we have:

$$\begin{aligned}&\bar{D}^B(\boldsymbol{R}(t)\text{: } R_j(t) = 1, R_k(t) = 0 \text{ for } k \neq j) = \frac{1}{N}\cdot \\ &\left(\bar{D}_j^B(R_j(t) = 1) + \sum_{k=1,k\neq j}^{N} \bar{D}_k^B(R_k(t) = 0)\right).\end{aligned} \tag{19}$$

Subtracting (19) from (18), we have:

$$\begin{aligned}(18) - (19) = \frac{1}{N}\Big(&\big(\bar{D}_j^B(R_j(t) = 0) - \bar{D}_j^B(R_j(t) = 1)\big) \\ &- \big(\bar{D}_i^B(R_i(t) = 0) - \bar{D}_i^B(R_i(t) = 1)\big)\Big).\end{aligned} \tag{20}$$

Clearly that if (20) is positive, transmitting source i leads to a greater $\bar{D}^B$ than transmitting source j. So we should transmit source j. Otherwise, we should transmit source i.

[10]Heh is the ancient Egyptian god of time.

By the $\bar{D}_i^B$ definition in (8), (20) becomes:

$$\lim_{T\to\infty}\frac{1}{N\cdot T}\cdot\sum_{\tau=1}^{T}\Big(D_j^B(\tau|R_j(t)=0)- D_j^B(\tau|R_j(t)=1)\Big)-\lim_{T\to\infty}\frac{1}{N\cdot T}\cdot\sum_{\tau=1}^{T}\Big(D_i^B(\tau|R_i(t)=0)-D_i^B(\tau|R_i(t)=1)\Big). \quad (21)$$

To decide whether source i or source j should transmit, we can ignore N, and focus solely on the following:

$$\begin{aligned}
&\lim_{T\to\infty}\frac{1}{T}\sum_{\tau=1}^{T}\Big(D_j^B(\tau|R_j(t)=0)-D_j^B(\tau|R_j(t)=1)\Big)\\
-\quad&\lim_{T\to\infty}\frac{1}{T}\sum_{\tau=1}^{T}\Big(D_i^B(\tau|R_i(t)=0)-D_i^B(\tau|R_i(t)=1)\Big)\\
=\quad&\lim_{T\to\infty}\frac{1}{T}\sum_{\tau=t+1}^{T}\Big(D_j^B(\tau|R_j(t)=0)-D_j^B(\tau|R_j(t)=1)\Big)\\
-\quad&\lim_{T\to\infty}\frac{1}{T}\sum_{\tau=t+1}^{T}\Big(D_i^B(\tau|R_i(t)=0)-D_i^B(\tau|R_i(t)=1)\Big).
\end{aligned} \quad (22)$$

To decide which node (i or j) has a higher priority, we want to know whether (22) is positive or not. This motivates us to define a *scheduling priority metric* $P_i(t)$ for each source $i=1,2,\cdots,N$ at each time t as follows:

$$P_i(t)=\lim_{T\to\infty}\frac{1}{T}\cdot\Big(\sum_{\tau=t+1}^{T}D_i^B(\tau|R_i(t)=0)-\sum_{\tau=t+1}^{T}D_i^B(\tau|R_i(t)=1)\Big). \quad (23)$$

From (22), at time t, if $P_i(t)\geq P_j(t)$ for any $j\neq i$, then we should transmit source i as its transmission minimizes $\bar{D}^B$ (as compared to any other nodes), i.e., we should select the source with the largest scheduling priority metric for transmission.

Now the remaining question is: How to calculate $P_i(t)$ for each source i according to (23)? Note that using (23) to calculate $P_i(t)$ is not possible for an online algorithm (no knowledge of the future). Nevertheless, we can use the scheduling priority metric in (23) as a guidance to develop an online algorithm. Specifically, we propose to estimate $P_i(t)$ based on past history (without any knowledge of the future) in the next section.

B. Estimating Priority Metric $P_i(t)$

Denote $\tilde{P}_i(t)$ as an estimate of $P_i(t)$ in (23). The goal of this section is to develop an expression for $\tilde{P}_i(t)$.

We perform this calculation by conditioning on the status of $D_i^B(\tau)$ at time $\tau=t$, i.e., $D_i^B(t)>0$ or $D_i^B(t)=0$.

Case 1: $D_i^B(t)>0$. In this case, to calculate $\tilde{P}_i(t)$, we first estimate $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=0)$, and then estimate $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=1)$.

Estimating $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=0)$: When $D_i^B(t)>0$ and $R_i(t)=0$, starting from time $\tau=t+1$, $D_i^B(\tau)$ increases linearly with time τ. Hence we estimate $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=0)$ as follows:

$$\begin{aligned}
&\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=0)\\
&\overset{\text{estimated by}}{=}\sum_{k=1}^{T-t}(D_i^B(t)+k)\\
&=(T-t)\cdot D_i^B(t)+\frac{1}{2}\cdot(T-t)\cdot(T-t+1).
\end{aligned} \quad (24)$$

Estimating $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=1)$: When $D_i^B(t)>0$ and $R_i(t)=1$, starting from time $\tau=t+1$, $D_i^B(\tau)$ first remains 0 for a certain number of time slots and then increase linearly with time. Denote $m_i(t)$ as the number of time slots for $D_i^B(\tau)$ to remain 0, i.e., from time $\tau=t+1$ to time $\tau=t+m_i(t)$, $D_i^B(\tau)=0$, and starting from time $\tau=t+m_i(t)+1$, $D_i^B(\tau)$ increases linearly with time. Hence we estimate $\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=1)$ as follows:

$$\begin{aligned}
&\sum_{\tau=t+1}^{T}D_i^B(\tau|D_i^B(t)>0,R_i(t)=1)\\
&\overset{\text{estimated by}}{=}\sum_{k=1}^{T-t-m_i(t)}k\\
&=\frac{1}{2}\cdot(T-t-m_i(t))\cdot(T-t-m_i(t)+1).
\end{aligned} \quad (25)$$

Now the remaining question is: How to calculate $m_i(t)$? The result is stated in the following lemma.

Lemma 3: For $m_i(t)$, we have:

$$\begin{aligned}
\mathbb{E}[m_i(t)|c_i^S(t-1)=0]&=\frac{2-p_i}{1-p_i}-(p_i+q_i);\\
\mathbb{E}[m_i(t)|c_i^S(t-1)=1]&=q_i+\frac{1-q_i^2}{1-p_i}.
\end{aligned}$$

We offer a proof sketch here. A complete proof is given in Appendix B. Based on source i's Markov model in Fig. 2, it can be shown that (i) $\frac{2-p_i}{1-p_i}-(p_i+q_i)$ is the expected number of time slots when $D_i^B(\tau)$ remains 0 since $\tau=t+1$, given that $c_i^S(t-1)=0$ and $R_i(t)=1$; and $q_i+\frac{1-q_i^2}{1-p_i}$ is the expected number of time slots when $D_i^B(\tau)$ remains 0 since $\tau=t+1$, given that $c_i^S(t-1)=1$ and $R_i(t)=1$.

Considering that

$$\tilde{P}_i(t|D_i^B(t)>0)=\lim_{T\to\infty}\frac{1}{T}\cdot\big((24)-(25)\big)=D_i^B(t)+m_i(t),$$

after we use $\mathbb{E}[m_i(t)]$ as an estimate for $m_i(t)$ in (25), we can find $\tilde{P}_i(t|D_i^B(t)>0)$:

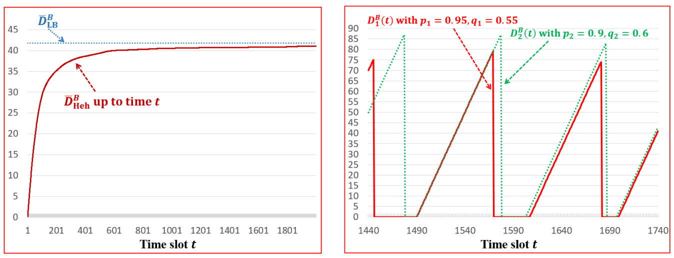
$$\tilde{P}_i(t|D_i^B(t)>0,c_i^S(t-1)=0)=D_i^B(t)+\frac{2-p_i}{1-p_i}-(p_i+q_i);$$

and

$$\tilde{P}_i(t|D_i^B(t)>0,c_i^S(t-1)=1)=D_i^B(t)+q_i+\frac{1-q_i^2}{1-p_i}.$$

Algorithm 1 Heh's Scheduling at Time t at the BS

1: Calculate $\tilde{P}_i(t)$ according to (26) for all i.
2: Transmit one source which has the largest $\tilde{P}_i(t)$.



(a) $\bar{D}^B_{\text{Heh}}$ vs $\bar{D}^B_{\text{LB}}$.

(b) $D^B_i(t)$ of two sources.

Fig. 3. Behavior of $\bar{D}^B_{\text{Heh}}$ and $D^B_i(t)$ in a case study ($N = 100$).

Case 2: $D^B_i(t) = 0$. In this case, to calculate $\tilde{P}_i(t)$, we first estimate $\sum_{\tau=t+1}^{T} D^B_i(\tau|D^B_i(t) = 0, R_i(t) = 0)$, and then estimate $\sum_{\tau=t+1}^{T} D^B_i(\tau|D^B_i(t) = 0, R_i(t) = 1)$. Following an analysis similar to our Case 1, we can find $\tilde{P}_i(t|D^B_i(t) = 0)$:

$$\tilde{P}_i(t|D^B_i(t) = 0, c^S_i(t-1) = 0) = 2 - (p_i + q_i);$$

and

$$\tilde{P}_i(t|D^B_i(t) = 0, c^S_i(t-1) = 1) = q_i + q_i \cdot \frac{1 - q_i}{1 - p_i}.$$

We summarize our results for $\tilde{P}_i(t)$ as follows:

$$\tilde{P}_i(t) = \begin{cases} D^B_i(t) + \frac{2-p_i}{1-p_i} - (p_i + q_i) & \text{if } D^B_i(t) > 0 \text{ and } c^S_i(t-1) = 0; \\ D^B_i(t) + q_i + \frac{1-q_i^2}{1-p_i} & \text{if } D^B_i(t) > 0 \text{ and } c^S_i(t-1) = 1; \\ 2 - (p_i + q_i) & \text{if } D^B_i(t) = 0 \text{ and } c^S_i(t-1) = 0; \\ q_i + q_i \cdot \frac{1-q_i}{1-p_i} & \text{if } D^B_i(t) = 0 \text{ and } c^S_i(t-1) = 1. \end{cases} \quad (26)$$

Note that in (26), the calculation of $\tilde{P}_i(t)$ at time t only requires the knowledge of $D^B_i(t)$, $c^S_i(t-1)$, p_i, and q_i, all of which do not require any information from the future.

C. Summary of Algorithm—Heh

By calculating the estimated scheduling priority metric $\tilde{P}_i(t)$ for each source i per (26), our scheduler Heh selects one source with the largest $\tilde{P}_i(t)$ for transmission at each time t. Details of Heh are given in Algorithm 1.

Now we discuss the time complexity of Heh. To calculate $\tilde{P}_i(t)$ for $i = 1, 2, \cdots, N$, the time complexity is $O(N)$. After that, to select the largest $\tilde{P}_i(t)$, the time complexity is $O(N)$. Thus, the total time complexity of Heh at each time t is $O(N)$.

V. Performance Evaluation

In this section we evaluate the performance of Heh. First, we use a case study to demonstrate how $\bar{D}^B$ and $D^B_i(t)$ achieved by Heh evolve with time t, respectively, and compare it to the lower bound $\bar{D}^B_{\text{LB}}$ that we derived in Lemma 2. Then we investigate how the $\bar{D}^B$ achieved by Heh is affected by different system parameters. For simplicity, we denote $\bar{D}^B_{\text{Heh}}$ as the $\bar{D}^B$ achieved by Heh. For each instance in our simulation, we simulate $T = 1,000,000$ time slots.

A. A Case Study

We set $N = 100$, and for each source i, both p_i and q_i are randomly generated by a uniform distribution from $[0.5, 0.99]$ (p_i and q_i can be different). Using the lower bound expression in (16), we find that $\bar{D}^B_{\text{LB}} = 41.77$. Besides, we find that $\bar{D}^B_{\text{Heh}}$ (when $t = T$) $= 41.84$, which is very close to our lower bound $\bar{D}^B_{\text{LB}}$. Moreover, in Fig. 3(a), we plot $\bar{D}^B_{\text{Heh}}$ from time $t = 1$ to time $t = 2000$. We observe that $\bar{D}^B_{\text{Heh}}$ up to time t converges quickly to $\bar{D}^B_{\text{LB}}$ and $\bar{D}^B_{\text{Heh}}$ (at time T). Since the minimum $\bar{D}^B_{\text{OPT}}$ lies between $\bar{D}^B_{\text{Heh}}$ and $\bar{D}^B_{\text{LB}}$, we conclude that:

- The lower bound $\bar{D}^B_{\text{LB}}$ is very tight;
- $\bar{D}^B_{\text{Heh}}$ is near-optimal.

To see the behavior of $D^B_i(t)$ for each source i, we choose sources $i = 1$ and $i = 2$ and plot $D^B_1(t)$ and $D^B_2(t)$ in Fig. 3(b) from time $t = 1440$ to time $t = 1740$. For source 1, $p_1 = 0.95$ and $p_2 = 0.55$. For source 2, $p_1 = 0.9$ and $p_2 = 0.6$. For each of the two sources, $D^B_i(t)$ remains 0 for multiple time slots and then increases linearly with time after each transmission.

B. Varying Number of Sources

We evaluate Heh under varying number of source nodes (N). For each source $i = 1, 2, \cdots, N$, p_i and q_i are generated in the same way as in the case study. In Table I, we present results of normalized $\bar{D}^B_{\text{Heh}}$, i.e., $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$, for N varying from 50 to 150. For each N, we simulate 100 instances (each with different p_i's and q_i's). For all cases, the mean of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ is within 1.07 and the variance is within 0.001. Further the maximum of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ is under 1.16 while the minimum is 1.

C. Varying Transition Probabilities

Finally we consider the impact of p_i's and q_i's on Heh's performance. Instead of varying p_i's and q_i's completely random, we consider some scenarios that have real world implications.

- *Rare Event Reporting* Based on the Markov model in Fig. 2, this corresponds to the scenario of large p_i's and small q_i's, where each source tends to remain mostly in the 0 state and rarely makes a transition to the 1 state. Even it goes to 1 state, it tends to go back to the 0 state more than it would stay in the 1 state. So we generate p_i's and q_i's randomly following a uniform distribution from $[0.7, 0.99]$ and $q_i = 1 - p_i$, respectively. We consider $N = 50$ and $N = 100$, respectively. For each N, we simulate 100 instances (each with different p_i's and q_i's) and give the results of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ in Table II, where the mean of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ is within 1.02, the variance is 0.00, and the maximum is under 1.09.
- *Active Surveillance* This corresponds to the scenario of large q_i's and small p_i's in the Markov model (see Fig. 2). In this scenario, each source tends to remain mostly in the 1 state and rarely makes a transition to the 0 state. When it goes to 0 state, it tends to return to the 1 state more than it would stay in the 0 state. So we generate q_i's

TABLE I
THE RATIO $\bar{D}^B_{\text{HEH}}/\bar{D}^B_{\text{LB}}$ UNDER DIFFERENT N ($p_i \in [0.5, 0.99]$ AND $q_i \in [0.5, 0.99]$).

N	50	60	70	80	90	100	110	120	130	140	150
Mean	1.063	1.042	1.031	1.026	1.017	1.015	1.011	1.009	1.005	1.003	1.002
Variance	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Min	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Max	1.158	1.109	1.101	1.071	1.059	1.052	1.041	1.039	1.038	1.028	1.027

TABLE II
THE RATIO $\bar{D}^B_{\text{HEH}}/\bar{D}^B_{\text{LB}}$ IN DIFFERENT SCENARIOS.

Scenario	Mean	Variance	Max
$N = 50, p_i \in [0.7, 0.99], q_i = 1 - p_i$	1.013	0.000	1.084
$N = 50, q_i \in [0.7, 0.99], p_i = 1 - q_i$	1.002	0.000	1.003
$N = 100, p_i \in [0.7, 0.99], q_i = 1 - p_i$	1.004	0.000	1.042
$N = 100, q_i \in [0.7, 0.99], p_i = 1 - q_i$	1.001	0.000	1.001

and p_i's randomly following a uniform distribution from $[0.7, 0.99]$ and $p_i = 1 - q_i$, respectively. We consider $N = 50$ and $N = 100$, respectively. For each N, we simulate 100 instances (each with different p_i's and q_i's) and also give the results of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ in Table II. We observe that the mean of $\bar{D}^B_{\text{Heh}}/\bar{D}^B_{\text{LB}}$ is within 1.01, the variance is 0.00, and the maximum is 1.01.

In summary, all of our results in this section indicate that $\bar{D}^B_{\text{Heh}}$ obtained by Heh is near-optimal. Further, the lower bound that we derived in Section III is very tight.

VI. CONCLUSIONS

In this paper, we introduced a metric called Ao^{2}I as an improvement of state-of-the-art metric AoII. Ao^{2}I quantifies the time lapse since the first time instance when stored information has been outdated at its source. We investigated a scheduling problem for minimizing Ao^{2}I in an IoT data collection network. We derived a theoretical lower bound for the minimum Ao^{2}I that can be used as a benchmark for any scheduler. Then we presented Heh—a low-complexity online scheduler to minimize Ao^{2}I. The design of Heh was based on the estimation of a novel offline scheduling priority metric in the absence of knowledge of the future. Through extensive simulations, we showed that Heh is near-optimal for minimizing Ao^{2}I and the lower bound is very tight.

ACKNOWLEDGMENT

This research was supported in part by ONR under MURI Grant N00014-19-1-2621, Virginia Commonwealth Cyber Initiative (CCI), and Virginia Tech Institute for Critical Technology and Applied Science (ICTAS).

APPENDIX

A. Proof of Lemma 1

Proof: We prove that for any scheduler that achieves $\bar{\boldsymbol{S}}$, the following must hold for its $\bar{D}^B_i$ for each $i = 1, 2, \cdots, N$:

$$\bar{D}^B_i \geq \max\left\{0,\ \frac{1}{2}\cdot\left(\frac{1}{\bar{S}_i} - (\bar{S}_i + 2)\cdot r_i + 1\right)\right\}. \tag{27}$$

For each $i = 1, 2, \cdots, N$, it is straightforward that $\bar{D}^B_i \geq 0$. So in this proof we focus on proving

$$\bar{D}^B_i \ \geq\ \frac{1}{2}\cdot\left(\frac{1}{\bar{S}_i} - (\bar{S}_i + 2)\cdot r_i + 1\right). \tag{28}$$

Denote $M_i(T)$ as the total number of samples transmitted by source i for the time horizon from time 1 to time T, i.e.,

$$M_i(T) \ =\ \sum_{t=1}^{T} S_i(t). \tag{29}$$

Denote $I_i(k)$ as the number of slots between the $(k-1)$-th and k-th sample transmissions from source i. After the last sample transmission from source i, the number of remaining slots is R_i. It is clear that for each i we have

$$T \ =\ \sum_{k=1}^{M_i(T)} I_i(k) + R_i. \tag{30}$$

For simplicity, suppose $D^B_i(1) = 0$ for all i. Note that as $\bar{D}^B_i$ is defined for $T \to \infty$, $\bar{D}^B_i$ does not depend on the value of $D^B_i(1)$. By the $D^B_i(t)$ definition, from the time which follows the $(k-1)$-th sample transmission to the time when the k-th sample is transmitted, $D^B_i(t)$ evolves as $\{0, 0, \cdots, 0, 1, 2, \cdots\}$. The reason why $D^B_i(t)$ remains 0 for certain period of time slots after transmitting the $(k-1)$-th sample is because the sample content of source i remains unchanged. We denote $u_i(k)$ $(1 \leq u_i(k) \leq I_i(k))$ as the number of slots when $D^B_i(t) = 0$ after source i transmits its $(k-1)$-th sample, and hence $D^B_i(t)$ evolves as $\{0, 0, \cdots, 0, 1, 2, \cdots, I_i(k) - u_i(k)\}$. Similarly, we denote $u_i(M_i(T)+1)$ $(1 \leq u_i(M_i(T)+1) \leq R_i)$ as the number of slots when $D^B_i(t) = 0$ after source i transmits the $M_i(T)$-th sample. As a result, the following holds based on (30):

$$\begin{aligned}&\tfrac{1}{T}\cdot\textstyle\sum_{t=1}^{T} D^B_i(t) = \\ &\tfrac{1}{2}\cdot\Big(\tfrac{M_i(T)}{T}\cdot\tfrac{1}{M_i(T)}\cdot\textstyle\sum_{k=1}^{M_i(T)}(I_i(k) - u_i(k))^2 + \\ &\tfrac{(R_i - u_i(M_i(T)+1))^2}{T} - \tfrac{1}{T}\textstyle\sum_{k=1}^{M_i(T)+1} u_i(k) + 1\Big).\end{aligned} \tag{31}$$

Define the operator $\bar{\mathbb{M}}(\cdot)$ to compute the sample mean, e.g.,

$$\bar{\mathbb{M}}(I_i) \ =\ \frac{1}{M_i(T)}\cdot\sum_{k=1}^{M_i(T)} I_i(k). \tag{32}$$

Then applying Jensen's inequality to (31) leads to

$$\frac{1}{T}\cdot\sum_{t=1}^{T}D_i^B(t) \;\geq\; \frac{1}{2}\cdot\Bigg(\frac{M_i(T)}{T}\cdot\big(\bar{\mathbb{M}}(I_i-u_i)\big)^2+ \frac{(R_i-u_i(M_i(T)+1))^2}{T}-\frac{1}{T}\sum_{k=1}^{M_i(T)+1}u_i(k)+1\Bigg)$$
$$=\frac{1}{2}\cdot\Bigg(\frac{M_i(T)}{T}\cdot\Big(\big(\bar{\mathbb{M}}(I_i)-\bar{\mathbb{M}}(u_i)\big)^2-\bar{\mathbb{M}}(u_i)\Big)+ \frac{(R_i-u_i(M_i(T)+1))^2-u_i(M_i(T)+1)}{T}+1\Bigg). \tag{33}$$

Combining (30) into (32) gives

$$\frac{1}{T}\cdot\sum_{t=1}^{T}D_i^B(t) \;\geq\; \frac{1}{2}\cdot\Bigg(\frac{(T-R_i)^2}{T\cdot M_i(T)}+\frac{M_i(T)}{T}\cdot\big(\bar{\mathbb{M}}(u_i)\big)^2 -\frac{2\cdot\bar{\mathbb{M}}(u_i)}{T}\cdot(T-R_i)-\frac{M_i(T)\cdot\bar{\mathbb{M}}(u_i)}{T}+ \frac{(R_i-u_i(M_i(T)+1))^2-u_i(M_i(T)+1)}{T}+1\Bigg). \tag{34}$$

By minimizing the RHS of the above inequality analytically with respect to the variable R_i, we have

$$\frac{1}{T}\cdot\sum_{t=1}^{T}D_i^B(t) \;\geq\; \frac{1}{2}\cdot\Bigg(\frac{T}{1+M_i(T)}+ \frac{M_i(T)}{T}\cdot\Big(\big(\bar{\mathbb{M}}(u_i)\big)^2-\bar{\mathbb{M}}(u_i)\Big)-2\cdot\bar{\mathbb{M}}(u_i)+ \frac{(u_i(M_i(T)+1))^2-u_i(M_i(T)+1)}{T}+\frac{2\cdot(\bar{\mathbb{M}}(u_i)-u_i(M_i(T)+1))}{1+M_i(T)}- \frac{M_i(T)\cdot(\bar{\mathbb{M}}(u_i)-u_i(M_i(T)+1))^2}{T\cdot(1+M_i(T))}+1\Bigg). \tag{35}$$

Due to Jensen's inequality we have

$$\mathbb{E}\left[\frac{T}{1+M_i(T)}\right]\geq\frac{1}{\frac{1}{T}+\mathbb{E}\left[\frac{M_i(T)}{T}\right]}. \tag{36}$$

Based on the definition of $M_i(T)$ we have

$$\lim_{T\to\infty}\mathbb{E}\left[\frac{M_i(T)}{T}\right] \;=\; \bar{S}_i. \tag{37}$$

For each $k=1,2,\cdots,M_i(T)+1$, we have

$$0 \;\leq\; \mathbb{E}[u_i(k)] \;\leq\; r_i. \tag{38}$$

The inequality in (38) holds due to the following: If source i is in state 0 when we transmit its $(k-1)$-th sample, we have:

$$\mathbb{P}\{u_i(k)=m\} \;=\; (1-p_i)\cdot p_i^{m-1} \text{ for } m\geq 1.$$

Otherwise if source i is in state 1 when we transmit its $(k-1)$-th sample, we have $\mathbb{P}\{u_i(k)=1\}=q_i$ and

$$\mathbb{P}\{u_i(k)=m\}=(1-q_i)\cdot(1-p_i)\cdot p_i^{m-2} \text{ for } m\geq 2.$$

Hence we have

$$\mathbb{E}[u_i(k)]\leq\sum_{m=1}^{\infty}m\cdot\mathbb{P}\{u_i(k)=m\}\leq r_i.$$

Considering that $\bar{\mathbb{M}}(u_i)\leq\max_k u_i(k)$, we have

$$0 \;\leq\; \mathbb{E}[\bar{\mathbb{M}}(u_i)] \;\leq\; r_i. \tag{39}$$

Moreover, we have the following

$$\lim_{T\to\infty}\mathbb{E}\left[\frac{\sum_{k=1}^{M_i(T)}u_i(k)}{T}\right]\leq\bar{S}_i\cdot r_i, \tag{40}$$

and

$$\lim_{T\to\infty}M_i(T)=\infty,\;\;\lim_{T\to\infty}\frac{T\cdot(1+M_i(T))}{M_i(T)}=\infty. \tag{41}$$

Combining (36), (37), (38), (39), (40), (41) with (35), we can prove that (28) is true. Hence, (27) holds. ■

B. Proof of Lemma 3

Proof: Let us define $v_i(t)$ as the number of time slots for $D_i^B(\tau)=0$ since the time $\tau=t+1$. In this proof, we prove

$$\mathbb{E}[v_i(t)|c_i^S(t-1)=0,R_i(t)=1]=\frac{2-p_i}{1-p_i}-(p_i+q_i); \tag{42}$$

$$\mathbb{E}[v_i(t)|c_i^S(t-1)=1,R_i(t)=1]=q_i+\frac{1-q_i^2}{1-p_i}. \tag{43}$$

It is clear this lemma will hold if we can prove (42) and (43). In order to prove (42) and (43), we first prove:

$$\mathbb{E}[v_i(t)|c_i^S(t-1)=0,D_i^B(t)=0,R_i(t)=0]=\frac{p_i}{1-p_i}; \tag{44}$$

$$\mathbb{E}[v_i(t)|c_i^S(t-1)=1,D_i^B(t)=0,R_i(t)=0]=\frac{1-q_i}{1-p_i}. \tag{45}$$

Consider $D_i^B(t)=0$ and $R_i(t)=0$. In this case $v_i(t)=m$ indicates $D_i^B(t+1)=D_i^B(t+2)=\cdots=D_i^B(t+m)=0$ and $D_i^B(t+m+1)=1$, which requires $c_i^S(t)=c_i^S(t+1)=\cdots=c_i^S(t+m-1)=0$ and $c_i^S(t+m)=1$.

From source i's Markov model, if $c_i^S(t-1)=0$, we have

$$\mathbb{P}\{v_i(t)=m|c_i^S(t-1)=0,D_i^B(t)=0,R_i(t)=0\} =(1-p_i)\cdot p_i^m \text{ for any } m\geq 1. \tag{46}$$

(44) holds directly due to (46).

From source i's Markov model, if $c_i^S(t-1)=1$, we have

$$\mathbb{P}\{v_i(t)=m|c_i^S(t-1)=1,D_i^B(t)=0,R_i(t)=0\} =(1-q_i)\cdot(1-p_i)\cdot p_i^{m-1} \text{ for any } m\geq 1. \tag{47}$$

(45) holds directly due to (47).

Now with (44) and (45), if $c_i^S(t-1)=0$, we have

$$\begin{aligned}&\mathbb{E}[v_i(t)|c_i^S(t-1)=0,R_i(t)=1]=1+\\&\mathbb{E}[v_i(t+1)|c_i^S(t)=0,D_i^B(t+1)=0,R_i(t+1)=0]\cdot p_i\\&+\mathbb{E}[v_i(t+1)|c_i^S(t)=1,D_i^B(t+1)=0,R_i(t+1)=0]\\&\cdot(1-p_i)=1+\frac{1}{1-p_i}-(p_i+q_i);\end{aligned} \tag{48}$$

Similarly, if $c_i^S(t-1)=1$, we can prove that

$$\mathbb{E}[v_i(t)|c_i^S(t-1)=1,R_i(t)=1]=q_i+\frac{1-q_i^2}{1-p_i}, \tag{49}$$

which completes the proof. ■

References

[1] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. of IEEE SECON*, pp. 350–358, Salt Lake City, UT, USA, June 27–30, 2011.

[2] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *Proc. of IEEE INFOCOM*, pp. 2731–2735, Orlando, FL, USA, March 25–30, 2012.

[3] Y. Sun, "A collection of recent papers on the age of information," [Online]. Available: http://webhome.auburn.edu/~yzs0078/. [Accessed: January 12, 2022].

[4] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Age-optimal sampling and transmission scheduling in multi-source systems," in *Proc. of ACM MobiHoc*, pp. 121–130, Catania, Italy, July 2–5, 2019.

[5] I. Kadota and E. Modiano, "Minimizing the age of information in wireless networks with stochastic arrivals," in *Proc. of ACM MobiHoc*, pp. 1173–1185, Catania, Italy, July 2–5, 2019.

[6] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. of IEEE ISIT*, pp. 2634–2638, Vail, CO, USA, June 17–22, 2018.

[7] C. Li, S. Li, Y. Chen, T. Hou, and W. Lou, "Minimizing age of information under general models for IoT data collection," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 2256–2270, 2019.

[8] C. Li, Y. Huang, Y. Chen, B. Jalaian, Y. T. Hou, and W. Lou, "Kronos: A 5G scheduler for AoI minimization under dynamic channel conditions," in *Proc. of IEEE ICDCS*, pp. 1466–1475, Dallas, TX, USA, July 7–9, 2019.

[9] B. T. Bacinoglu and E. Uysal-Biyikoglu, "Scheduling status updates to minimize age of information with an energy harvesting sensor," in *Proc. of IEEE ISIT*, pp. 1122–1126, Aachen, Germany, June 25–30, 2017.

[10] Z. Zhou, C. Fu, C. J. Xue, and S. Han, "Transmit or discard: Optimizing data freshness in networked embedded systems with energy harvesting sources," in *Proc. of DAC*, 6 pages, Las Vegas, NV, USA, June 3–5, 2019.

[11] S. Feng and J. Yang, "Minimizing age of information for an energy harvesting source with updating failures," in *Proc. of IEEE ISIT*, pp. 2431–2435, Vail, CO, USA, June 17–22, 2018.

[12] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal transmission for energy harvesting sensors with finite batteries: Online policies," *IEEE Transactions on Information Theory*, vol. 66, no. 1, pp. 534–556, 2019.

[13] I. Kadota, A. Sinha, and E. Modiano, "Optimizing age of information in wireless networks with throughput constraints," in *Proc. of IEEE INFOCOM*, pp. 1844–1852, Honolulu, HI, USA, April 15–19, 2018.

[14] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proc. of ACM MobiHoc*, pp. 191–200, Los Angeles, CA, USA, June 25–28, 2018.

[15] Q. Liu, H. Zeng, and M. Chen, "Minimizing age-of-information with throughput requirements in multi-path network communication," in *Proc. of ACM MobiHoc*, pp. 41–50, Catania, Italy, July 2–5, 2019.

[16] I. Kadota, A. Sinha, and E. Modiano, "Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints," *IEEE/ACM Transactions on Networking*, vol. 27, no. 4, pp. 1359–1372, 2019.

[17] X. Zheng, S. Zhou, Z. Jiang, and Z. Niu, "Closed-form analysis of non-linear age of information in status updates with an energy harvesting transmitter," *IEEE Transactions on Wireless Communications*, vol. 18, no. 8, pp. 4129–4142, 2019.

[18] R. Talak and E. Modiano, "Age-delay tradeoffs in single server systems," in *Proc. of IEEE ISIT*, pp. 340–344, Paris, France, July 7–12, 2019.

[19] A. Maatouk, M. Assaad, and A. Ephremides, "On the age of information in a CSMA environment," *IEEE/ACM Transactions on Networking*, vol. 28, no. 2, pp. 818–831, 2020.

[20] C. Li, S. Li, Y. Chen, Y. T. Hou, and W. Lou, "AoI scheduling with maximum thresholds," in *Proc. of IEEE INFOCOM*, pp. 436–445, Virtual Conference, July 6–9, 2020.

[21] C. Li, Q. Liu, S. Li, Y. Chen, Y. T. Hou, and W. Lou, "On scheduling with AoI violation tolerance," in *Proc. of IEEE INFOCOM*, 10 pages, Virtual Conference, May 10–13, 2021.

[22] Q. Liu, C. Li, Y. T. Hou, W. Lou, and S. Kompella, "Aion: A bandwidth optimized scheduler with AoI guarantee," in *Proc. of IEEE INFOCOM*, 10 pages, Virtual Conference, May 10–13, 2021.

[23] G. D. Nguyen, S. Kompella, C. Kam, J. E. Wieselthier, and A. Ephremides, "Information freshness over an interference channel: A game theoretic view," in *Proc. of IEEE INFOCOM*, pp. 908–916, Honolulu, HI, USA, April 15–19, 2018.

[24] Z. Ning, P. Dong, X. Wang, X. Hu, L. Guo, B. Hu, Y. Guo, T. Qiu, and R. Kwok, "Mobile edge computing enabled 5G health monitoring for Internet of medical things: A decentralized game theoretic approach," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 2, pp. 463–478, 2020.

[25] G. D. Nguyen, S. Kompella, C. Kam, J. E. Wieselthier, and A. Ephremides, "Impact of hostile interference on information freshness: A game approach," in *Proc. of WiOpt*, 7 pages, Paris, France, May 15–19, 2017.

[26] S. Ioannidis, A. Chaintreau, and L. Massoulié, "Optimal and scalable distribution of content updates over a mobile social network," in *Proc. of IEEE INFOCOM*, pp. 1422–1430, Rio de Janeiro, Brazil, April 19–25, 2009.

[27] C. Kam, S. Kompella, and A. Ephremides, "Experimental evaluation of the age of information via emulation," in *Proc. of IEEE MILCOM*, pp. 1070–1075, Tampa, FL, USA, October 26–28, 2015.

[28] J. Zhong, R. D. Yates, and E. Soljanin, "Two freshness metrics for local cache refresh," in *Proc. of IEEE ISIT*, pp. 1924–1928, Vail, CO, USA, June 17–22, 2018.

[29] B. Yin, S. Zhang, Y. Cheng, L. X. Cai, Z. Jiang, S. Zhou, and Z. Niu, "Only those requested count: Proactive scheduling policies for minimizing effective age-of-information," in *Proc. of IEEE INFOCOM*, pp. 109–117, Paris, France, April 29–May 2, 2019.

[30] S. Feng and J. Yang, "Information freshness for timely detection of status changes," in *Proc. of IEEE CISS*, 6 pages, Princeton University, NJ, USA, March 18–20, 2020.

[31] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides, "The age of incorrect information: A new performance metric for status updates," *IEEE/ACM Transactions on Networking*, vol. 28, no. 5, pp. 2215–2228, 2020.

[32] Y. Chen and A. Ephremides, "Minimizing age of incorrect information for unreliable channel with power constraint," *arXiv preprint arXiv:2101.08908*, 2021.

[33] C. Kam, S. Kompella, and A. Ephremides, "Age of incorrect information for remote estimation of a binary markov source," in *Proc. of IEEE INFOCOM Workshops - Age of Information Workshop*, 6 pages, Virtual Conference, July 6–9, 2020.

[34] S. Kriouile and M. Assaad, "Minimizing the age of incorrect information for real-time tracking of markov remote sources," *arXiv preprint arXiv:2102.03245*, 2021.

[35] R. Talak, S. Karaman, and E. Modiano, "Optimizing information freshness in wireless networks under general interference constraints," in *Proc. of ACM MobiHoc*, pp. 61–70, Los Angeles, CA, USA, June 25–28, 2018.

[36] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *Proc. of IEEE ISIT*, pp. 321–325, Aachen, Germany, June 25–30, 2017.

[37] J. F. Kurose and K. W. Ross, "Computer networking: A top-down approach," Chapter 3, *Pearson Education*, 2017.

[38] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.2," [Online]. Available: http://cvxr.com/cvx. [Accessed: January 12, 2022].