Minimizing Age of Information under Arbitrary Arrival Model with Arbitrary Packet Size

[Extended Abstract]

Kumar Saurav  
School of Technology and Computer Science  
Tata Institute of Fundamental Research  
Mumbai, India  
kumar.saurav@tifr.res.in

Rahul Vaze  
School of Technology and Computer Science  
Tata Institute of Fundamental Research  
Mumbai, India  
rahul.vaze@gmail.com

1. INTRODUCTION
In networked systems such as internet-of-things, cyber-physical systems, etc., information timeliness is a critical requirement. Timely updates ensure that the information available at the subsystems (nodes) are accurate, and actions are well-coordinated. One popular metric to capture information timeliness is the age of information (AoI) [1]. At any time, for a source-destination pair, AoI is equal to the time elapsed since the generation time of the latest update of the source that has been completely transmitted to the destination. Thus, at time $t$, the AoI $\Delta(t) = t - \lambda(t)$, where $\lambda(t)$ denotes the generation time of the latest update of the source that has been completely transmitted until time $t$.

Note that unlike classical (packet-based) metrics such as flow-time, completion time, etc. that are defined for each update/packet, AoI is defined for a source, and at any time, only depends on the generation time of the latest completely transmitted update. Therefore, for minimizing the AoI, all updates need not be transmitted (it may be sufficient to transmit only a subset of updates). Thus, AoI minimization problems have a combinatorial feature, which makes them fundamentally different and analytically challenging compared to the classical packet-based scheduling problems.

Considered Problem. Consider a source-destination pair, where updates are generated at the source at arbitrary time instants $g_1, g_2, \ldots$, revealed causally. Moreover, for each update $i$, its size $s_i$ (transmission time required to completely transmit it) is also arbitrary, and revealed at the generation time $g_i$. At any time, to minimize the AoI, the source may transmit any one of the available updates, while incurring a transmission cost at rate $c$ for per unit time of transmission. The total cost incurred by the source (henceforth, the sum-average cost) is equal to the sum of the average AoI and the average transmission cost:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t (\Delta(\tau) + c \cdot \mathbb{1}(t))d\tau, \quad (1)$$

where $\mathbb{1}(t) = 1$ if an update is under transmission at time $t$, and 0 otherwise. The goal is to find a causal policy $\pi$ that minimizes the sum-average cost (1), where the possible decisions at any time are i) whether to preempt an ongoing update transmission on generation of a new update, and ii) if no update is under transmission, then choose which update to transmit (or not to transmit) among the available updates.

To characterize the performance of any causal policy under arbitrary arrival model, we use the metric of competitive ratio. Let $\mathcal{I} = \{(g_1, s_1), (g_2, s_2), \ldots\}$ denote a particular sequence of updates that are generated at the source. Moreover, let $\pi^*$ denote an optimal offline policy that knows $\mathcal{I}$ ahead of time. For any causal policy $\pi$, its competitive ratio is defined as the ratio of the cost (1) incurred by $\pi$ and $\pi^*$, respectively, maximized over all possible sequences $\mathcal{I}$.

Related Works. Competitive ratio analysis for the packet-based metrics such as flow-time, etc. is very well known. In particular, for a single source-destination pair, the SRPT (shortest remaining processing time) policy that always transmits the packet with the least remaining processing time (size), has a competitive ratio equal to 1 (i.e., SRPT is an optimal policy) [2]. For more involved models such as multiple servers, scheduling policies have been proposed whose competitive ratios have been bounded. Compared to the packet-based metrics, AoI is fundamentally different because it is source-based, and does not require transmitting all the updates generated at the source. Thus, an optimal offline policy $\pi^*$ which knows the generation times and sizes of all the updates in advance, can avoid transmitting large-sized updates, while a causal policy may not be able to do so. Therefore, for heuristic policies the competitive ratio can be large. For example, the SRPT policy that is optimal for minimizing the flow-time for a single source-destination pair also intuitively appears to be a reasonable policy for minimizing the AoI. However, note that if large number of updates are generated close to time $t = 0$ with marginally smaller size than other updates, then SRPT keeps transmitting these smaller updates for a long time instead of the recent updates. Therefore, for the AoI minimization problem, the competitive ratio of SRPT policy is arbitrarily large. Similar result can be shown for LCFS policy as well that at any time, transmits the most recently generated update, irrespective of size.

Another natural policy for AoI minimization is a greedy policy that at each time $t$, chooses to transmit that update for which the ratio of the reduction in AoI (upon its completion) and the remaining size at time $t$, is maximum. In the arbitrary arrival model, one can construct update generation sequences such that the greedy policy repeatedly
preempts the update under transmission without transmitting any update completely, for a long time. Recall that the AoI decreases only when an update is completely transmitted. Hence, establishing competitive ratio guarantees for the greedy policy is difficult.

As described above, both the SRPT and the greedy policy on their own may not have good competitive ratio performance. However, for certain classes of inputs (update generation sequences), individually, both have complementarily appealing features. With this motivation, we propose a policy called SRPT+, that at any point in time, either uses SRPT, or the greedy policy depending on the state of the system. We show that when the cost for transmission $c = 0$, the competitive ratio for SRPT+ is $CR_+ \leq 3$. Even when $c > 0$, by employing a simple threshold condition (3) (to restrict unnecessary transmission), we show that $CR_+ \leq 5$. Compared to prior works on AoI minimization (with and without transmission cost), this is the first result that provides a constant competitive ratio bound for the arbitrary arrival model.

2. CONTRIBUTIONS

SRPT+ Policy. Let $s_i(t)$ denote the remaining size of update $i$ at time $t$, and $\lambda(t)$ is the generation time of the latest update completely transmitted until time $t$. If update $i$ is transmitted starting at time $t$ (and not preempted ever), then its transmission completes at time $t + s_i(t)$, causing an instantaneous reduction in AoI of $\max\{g_i - \lambda(t), 0\}$ units at time instant $t + s_i(t)$. Thus, for each update $i$ available at time $t$, i.e., $s_i(t) > 0$, define the index

$$\gamma_i(t) = \frac{\max\{g_i - \lambda(t), 0\}}{s_i(t)},$$

which represents the ratio of the benefit obtained by transmitting update $i$, and the cost (time required to completely transmit it). Further, since there is cost for transmission and all updates need not be transmitted to minimize AoI, at any time $t$, we define $\Gamma(t)$ as the set of all remaining updates $i$ ($s_i(t) > 0$) such that the cost for their transmission is less compared to the benefit (reduction in average AoI) that would be obtained on their complete transmission. Formally, $\Gamma(t) = \{ i | s_i(t) > 0, \gamma_i(t) > 0, \text{and (3) is satisfied}\}$.

$$c_{s_i(t)} \leq (g_i - \lambda(t))\Delta(t) + (t - g_i)/2.$$  \hspace{1cm} (3)

Remark 1. In (3), $c_{s_i(t)}$ is the transmission cost for update $i$, while $(g_i - \lambda(t))$ is the reduction in AoI if the transmission of update $i$ completes. The term $\Delta(t) + (t - g_i)/2$ captures the fact that if the AoI has grown large, and no recent update has been generated with small size (since $g_i$), then update $i$ may be considered for transmission.

Consider the policy called SRPT+ (Algorithm 1) that at any time $t$, makes one of the following two decisions. i) If no update is under transmission, then it starts to transmit update $j^*(t) = \arg\max_{j \in \Gamma(t)} \gamma_j(t)$ (and idles if the set $\Gamma(t)$ is empty), and ii) if an update $i$ is under transmission, then update $i$ is preempted if and only if a new update $j$ is generated with size $s_j = s_i(t) \leq s_i(t)$.

Theorem 1. The proposed causal policy SRPT+ has a constant competitive ratio (denoted as $CR_+$). In particular,

1. for any transmission cost $c \geq 0$, we have $CR_+ \leq 5$.

Algorithm 1 SRPT+ Policy.

At any time $t$,

\begin{itemize}
  \item if an update $i$ is under transmission then
  \item if a new update $j$ is generated with size $s_j(t) \leq s_i(t)$ then
    \begin{itemize}
      \item preempt update $i$ and begin to transmit update $j$.
    \end{itemize}
  \item else
    \begin{itemize}
      \item continue transmitting update $i$.
    \end{itemize}
  \end{itemize}

\begin{itemize}
  \item if else if the set $\Gamma(t)$ is non-empty then
    \begin{itemize}
      \item begin to transmit update $j = \arg\max_{i \in \Gamma(t)} \gamma_i(t)$ (ties broken arbitrarily).
    \end{itemize}
  \item else
    \begin{itemize}
      \item idle until a new update is generated.
    \end{itemize}
\end{itemize}

2. when $c = 0$, we have $1.42 \leq CR_+ \leq 3$.

3. in the special case when the size of all the updates are equal, for any $c \geq 0$, we have $CR_+ \leq 2$.

Proof Idea. Recall that $g_i$ is the generation time of update $i$. For each update $i$, $r_i^\pi$ denote the earliest time instant when the transmission of an update $j$ with generation time $g_j \geq g_i$ completes under policy $\pi$. Define $\delta_i = g_i - g_{i-1}$ and $\nu_i^\pi = r_i^\pi - g_i$. To prove Theorem 1 we express the average AoI in terms of packet-based metrics $\delta_i$ and $\nu_i^\pi$ defined for each update $i$ generated at the source (see Figure 1). Then for each update $i$, we bound $\nu_i^\pi$ in terms of $\nu_i^\pi$, where the superscripts $+$ and $*$ denotes SRPT+ and an optimal offline policy $\pi^*$, respectively. Thus, we get a bound on the average AoI for SRPT+ in terms of the average AoI for $\pi^*$. The transmission cost for SRPT+ is bounded by comparing it against the incurred average AoI (by SRPT+).

Acknowledgments. This work was supported by the Department of Atomic Energy, Government of India, under project no. RTI4001.

3. REFERENCES

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

[2] L. Schrage. A proof of the optimality of the shortest remaining processing time discipline. Operations Research, 16(3):687–690, 1968.