Age-optimal Sampling and Transmission Scheduling in Multi-Source Systems

Ahmed M. Bedewy
Department of ECE
The Ohio State University
Columbus, OH
bedewy.2@osu.edu

Yin Sun
Department of ECE
Auburn University
Auburn, AL
yzs0078@auburn.edu

Sastry Kompella
Information Technology Division
Naval Research Laboratory,
Washington, DC
sk@ieee.org

Ness B. Shroff
Department of ECE and CSE
The Ohio State University
Columbus, OH
shroff.11@osu.edu

ABSTRACT

In this paper, we consider the problem of minimizing the *age of information* in a multi-source system, where sources communicate their update packets to a destination via a channel with random delay. Due to interference, only one source can be scheduled at a time. We consider the problem of finding a decision policy that controls the packet sampling times and schedules source transmissions to minimize the total average peak age (TaPA) and the total average age (TaA) of the sources. Our investigation of this problem results in an important *separation principle*: The optimal scheduling strategy and the optimal sampling strategy are independent from each other. In particular, we prove that, given the sampling times of the update packets, the Maximum Age First (MAF) scheduling strategy provides the best age performance among all scheduling strategies. This transforms our optimization problem into the optimal sampling problem, given that the decision policy follows the MAF scheduling strategy. Interestingly, we show that the zero-wait sampler (in which a packet is generated once the channel is idle) is optimal for minimizing the TaPA, while it does not always minimize the TaA. We use Dynamic Programming (DP) to investigate the optimal sampler for minimizing the TaA. Finally, we provide an approximate analysis of Bellman's equation to approximate the TaA-optimal sampler by a water-filling solution, which is quite close to the optimum in numerical evaluations.

KEYWORDS

Age of information; Data freshness; Sampling; Scheduling, Information update system; Multi-source

MobiHoc'19, July 2019, Catania, Italy

ACM ISBN 123-4567-24-567/08/06...$15.00
https://doi.org/10.475/123_4

ACM Reference format:
Ahmed M. Bedewy, Yin Sun, Sastry Kompella, and Ness B. Shroff. 2019. Age-optimal Sampling and Transmission Scheduling in Multi-Source Systems. In *Proceedings of ACM MobiHoc, Catania, Italy, July 2019 (MobiHoc'19),* 13 pages.
https://doi.org/10.475/123_4

1 INTRODUCTION

In recent years, significant attention has been paid to the *age of information* as a metric for data freshness. This is because there are a growing number of applications that require timely status updates in a variety of networked monitoring and control systems. Examples, include sensor and environment monitoring networks, surrounding monitoring autonomous vehicles, smart grid systems, etc. The age of information, or simply age, was introduced in [1, 9, 11, 20], and defined as the time elapsed since the most recently received update was generated. Unlike traditional packet-based metrics, such as throughput and delay, age is a destination-based metric that captures the information lag at the destination, and hence is more suitable to achieve the goal of timely updates.

Early studies characterized the age in many interesting variants of queueing models [10, 15, 19–22, 26, 38], in which the update packets arrive at the queue randomly as a Poisson process. Besides these queueing theoretical studies, the work in [3–6] showed that Last Generated First Served (LGFS)-type policies are (nearly) optimal for minimizing any non-decreasing functional of the age process in single flow multi-server and multi-hop networks. These results hold for general system settings that include arbitrary packet generation at the source and arbitrary packet arrival times at the transmitter queue. A generalization of these results was later considered in [30] for multi-flow multi-server queueing systems, under the condition that the packet generation and arrival times are synchronized across the flows.

Another line of research has considered the "generate-at-will" model [2, 31, 37], in which the generation time (sampling time) of the update packets are controllable. The work in [31] considered the optimal sampling problem for minimizing non-negative, non-decreasing age penalty functions, where a source can generate an update packet at any time and send it to the destination through

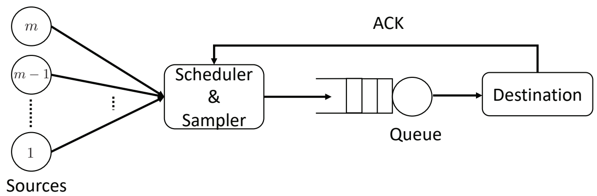


Figure 1: System model

a channel with random delay. Our work here is one step forward from the work in [31] by considering multi-source information update system as shown in Fig. 1, where sources send their update packets to the destination through a channel. Due to the resource limitation (only one source can send a packet through the channel at a time), a decision maker not only controls the packet generation times, but also schedules the sources. Thus, extending to the multi-source case is a challenging problem, as it requires seeking for the optimal policy that controls both the sampling and the scheduling to minimize the age.

The scheduling problem for multi-source networks with different scenarios was considered in [12–14, 16, 17, 32–35, 39]. In [12], the authors found that the scheduling problem for minimizing the age in wireless networks under physical interference constraints is NP-hard. Optimal scheduling for age minimization in a broadcast network was studied in [13, 14, 16, 17], where a sinlge source can be scheduled at a time. In contrast to our study, the generation of the update packets in [12–14, 16, 17] is uncontrollable and they arrive randomly at the transmitter. Age analysis of the status updates over a multiaccess channel was considered in [39]. The studies in [32–35] considered the age optimization problem in wireless network with general interference constraints and channel uncertainty. The considered sources in [32–35, 39] are active such that they can generate a new packet for each transmission (active sources are equivalent to zero-wait sampling strategy in our model, where a packet is generated from a source once this source is scheduled). Moreover, all the aforementioned studies for multi-source scheduling considered a time-slotted system, where a packet is transmitted in one time slot (i.e., a deterministic transmission time). Our investigation in this paper reveals that the zero-wait sampling strategy doesn't always minimize the age (the TaA in particular) in multi-source networks with random transmission times (which could be more than one time slot). Thus, our work here complements the studies in [12–14, 16, 17, 32–35, 39] by answering the following important question: What is the optimal policy that controls the packet generation times and the source scheduling to minimize the age in a multi-source information update system with random transmission times? To that end, the main contributions of this paper are outlined as follows:

- We formulate the problem of finding the optimal policy that controls the sampling and scheduling strategies to minimize two age of information metrics, namely the total average peak age (TaPA) and the total average age (TaA).
- We show that our optimization problem has an important *separation principle*: The optimal sampling strategy and the optimal scheduling strategy can be designed independently from each other. In particular, we use the stochastic ordering technique to show that, given the generation time of the update packets, the Maximum Age First (MAF) scheduling strategy provides a better age performance compared to any other scheduling strategy (Theorem 3.2). This *separation principle* helps us shrink our decision policy space and transform our complicated optimization problem into an optimal sampling problem for minimizing the TaPA and TaA by fixing the scheduling strategy to the MAF strategy.
- We formulate the optimal sampler problem for minimizing the TaPA. We show that the zero-wait sampler is the optimal sampler in this case (Theorem 3.3). However, interestingly, we find that the zero-wait sampler does not always minimize the TaA.
- We map the optimal sampling problem for minimizing the TaA into an equivalent optimization problem with a simpler form that enables us to use Dynamic Programming (DP) to obtain the optimal sampler. We show that there exists a stationary deterministic sampler that can achieve optimality (Theorem 3.5). Moreover, we show that the optimal sampler has a threshold property (Theorem 3.6) that helps in reducing the complexity of the relative value iteration (RVI) algorithm (by reducing the computations required along the system state space).
- Our optimal scheduling and sampling strategies can minimize the age for any random transmission times (with a discrete distribution and possibly more than one time slot), which cannot be addressed by prior studies (e.g. [12–14, 16, 17, 32–35, 39]). Although the optimality of the MAF scheduler was shown in [14, 17, 30], these studies either consider a time-slotted system [14, 17], or stochastic arrivals with exponential transmission times [30]. Our results here can be seen as an extension of these studies by generalizing the transmission times and controlling the packet generation times.
- Finally, in Section 4, we provide an approximate analysis of Bellman's equation which results in the water-filling solution as an approximate solution for the optimal sampler (for minimizing the TaA). The performance of the water-filling solution is evaluated in Fig. 5.

To the best of our knowledge, this is the first study that considers the optimization of the sampling and scheduling strategies to minimize the age in multi-source networks with random transmission times.

2 MODEL AND FORMULATION

2.1 Notations

For any random variable Z and an event A, let $\mathbb{E}[Z|A]$ denote the conditional expectation of Z for given A. We use $\mathbb{N}^+$ to represent the set of non-negative integers, $\mathbb{R}^+$ is the set of non-negative real numbers, $\mathbb{R}$ is the set of real numbers, and $\mathbb{R}^n$ is the set of n-dimensional real Euclidean space. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be two vectors in $\mathbb{R}^n$, then we denote $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for $i = 1, 2, \ldots, n$. Also, we use $x_{[i]}$ to denote the i-th largest component of vector $\mathbf{x}$. For any bounded set $\mathcal{X} \subset \mathbb{R}$, we use $\max \mathcal{X}$ to represent the maximum of set $\mathcal{X}$, i.e., $x^* = \max \mathcal{X}$ implies that $x \leq x^*$ for all $x \in \mathcal{X}$. A set $U \subseteq \mathbb{R}^n$ is called upper if $\mathbf{y} \in U$ whenever $\mathbf{y} \geq \mathbf{x}$ and $\mathbf{x} \in U$. We will need the following definitions:

Definition 2.1. **Univariate Stochastic Ordering:** [28] Let X and Y be two random variables. Then, X is said to be stochastically smaller than Y (denoted as $X \leq_{\text{st}} Y$), if

$$\mathbb{P}\{X > x\} \leq \mathbb{P}\{Y > x\}, \quad \forall x \in \mathbb{R}.$$

Definition 2.2. **Multivariate Stochastic Ordering:** [28] Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors. Then, $\mathbf{X}$ is said to be stochastically smaller than $\mathbf{Y}$ (denoted as $\mathbf{X} \leq_{\text{st}} \mathbf{Y}$), if

$$\mathbb{P}\{\mathbf{X} \in U\} \leq \mathbb{P}\{\mathbf{Y} \in U\}, \quad \text{for all upper sets} \quad U \subseteq \mathbb{R}^n.$$

Definition 2.3. **Stochastic Ordering of Stochastic Processes:** [28] Let $\{X(t), t \in [0,\infty)\}$ and $\{Y(t), t \in [0,\infty)\}$ be two stochastic processes. Then, $\{X(t), t \in [0,\infty)\}$ is said to be stochastically smaller than $\{Y(t), t \in [0,\infty)\}$ (denoted by $\{X(t), t \in [0,\infty)\} \leq_{\text{st}} \{Y(t), t \in [0,\infty)\}$), if, for all choices of an integer n and $t_1 < t_2 < \ldots < t_n$ in $[0,\infty)$, it holds that

$$(X(t_1), X(t_2), \ldots, X(t_n)) \leq_{\text{st}} (Y(t_1), Y(t_2), \ldots, Y(t_n)), \tag{1}$$

where the multivariate stochastic ordering in (1) was defined in Definition 2.2.

2.2 System Model

We consider a status update system with m sources as shown in Fig. 1, where each source observes a time-varying process. An update packet is generated from a source and is then sent over an error-free delay channel to the destination, where only one packet can be sent at a time. A decision maker (a controller) controls the generation time of the update packets and schedules source transmissions. This is known as "generate-at-will" model [2, 31, 37] (i.e., the decision maker can generate the update packets at any time). We use S_i to denote the generation time of the i-th generated packet, called packet i. The channel is modeled as First-Come First-Served (FCFS) queue with random *i.i.d.* service time Y_i, where Y_i represents the service time of packet i, $Y_i \in \mathcal{Y}$, and $\mathcal{Y} \subset \mathbb{R}^+$ is a finite and bounded set. We suppose that the decision maker knows the idle/busy state of the server through acknowledgments (ACKs) from the destination with zero delay. To avoid unnecessary waiting time in the queue, there is no need to generate an update packet during the busy periods. Let D_i denote the delivery time of packet i, where $D_i = S_i + Y_i$. After the delivery of packet i at time D_i, the decision maker may insert a waiting time Z_i before generating a new packet (hence, $S_{i+1} = D_i + Z_i$)[1], where $Z_i \in \mathcal{Z}$, and $\mathcal{Z} \subset \mathbb{R}^+$ is a finite and bounded set. We use M to represent the the maximum amount of waiting time allowed by the system, i.e., $M = \max \mathcal{Z}$.

We use $\mathcal{G}_l$ to represent the set of generation times of the update packets that are generated from source l. At any time t, the most recently delivered packet from source l is generated at time

$$U_l(t) = \max\{S_i : S_i \in \mathcal{G}_l, D_i \leq t\}. \tag{2}$$

The *age-of-information*, or simply the *age*, for source l is defined as [1, 9, 11, 20]

$$\Delta_l(t) = t - U_l(t). \tag{3}$$

[1]We suppose that $D_0 = 0$. Thus, we have $S_1 = Z_0$.

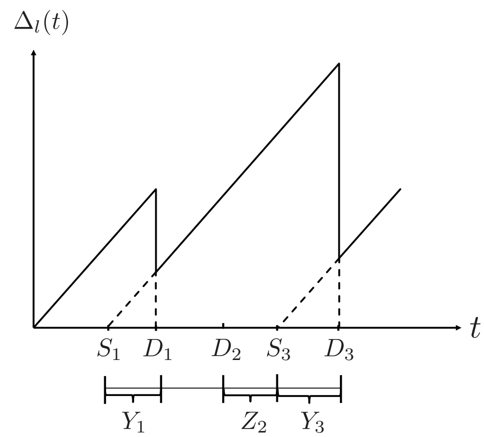


Figure 2: The age of source l ($\Delta_l(t)$), where we suppose that $S_1, S_3 \in \mathcal{G}_l$.

As shown in Fig. 2, the age increases linearly with t but is reset to a smaller value with the delivery of a fresher packet. The age process for source l is given by $\{\Delta_l(t), t \geq 0\}$. We suppose that the initial age values ($\Delta_l(0^-)$ for all l) are known to the system.

2.3 Decision Policies

A decision policy, denoted by d, controls the following: i) source scheduling, i.e., determines the source to be served at each transmission opportunity, ii) the packet generation times $(S_1, S_2, \ldots)$, or equivalently, the sequence of waiting times $(Z_0, Z_1, \ldots)$. Let $\mathcal{D}$ denote the set of causal decision policies in which decisions are made based on the history and current states of the system.

After each delivery, the decision maker chooses the source to be served, and imposes a waiting time before the generation of the new packet. Next, we present our optimization problems.

2.4 Optimization Problem

Let r_i represent the index of the source from which packet i is generated. We define two metrics to assess the long term age performance over our status update system in (4) and (5). Consider the time interval $[0, D_n]$, for any decision policy d, we define the total average peak age (TaPA) as

$$\Delta_{\text{peak}}(d) = \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^{n+1} \Delta_{r_i}(D_i^-)\right], \tag{4}$$

and the total average age per unit time (TaA) as

$$\Delta_{\text{avg}}(d) = \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{l=1}^{m} \int_0^{D_n} \Delta_l(t)dt\right]}{\mathbb{E}[D_n]}. \tag{5}$$

In this paper, we aim to minimize both the TaPA and the TaA separately. Thus, our optimization problems can be formulated as follows. We seek a decision policy d that solve the following optimization problems

$$\bar{r}_{\text{opt}} \triangleq \min_{d \in \mathcal{D}} \Delta_{\text{peak}}(d), \tag{6}$$

and

$$\bar{g}_{\text{opt}} \triangleq \min_{d\in\mathcal{D}} \Delta_{\text{avg}}(d), \tag{7}$$

where $\bar{r}_{\text{opt}}$ and $\bar{g}_{\text{opt}}$ are the optimum objective values of Problems (6) and (7), respectively. Due to the large decision policy space, our optimization problem is quite challenging. In particular, it is required to seek the optimal decision policy that controls both the sampling and scheduling strategies to minimize the TaPA and the TaA. In the next section, we discuss our approach to tackle this optimization problem.

3 OPTIMAL DECISION POLICY

We first show that our optimization problems in (6) and (7) have an important separation principle: The scheduling strategy and the sampling strategy can be designed independently from each other. To that end, we show that, given the generation time of the update packets, following the Maximum Age First (MAF) scheduling strategy provides the best age performance compared to following any other scheduling strategy. By this, it remains to address the question of finding the best sampling strategies that solve problems (6) and (7). Next, we present our approach to solve our optimization problems in detail.

3.1 Optimal Scheduling Strategy

We start by defining the MAF scheduling strategy as follows:

Definition 3.1. Maximum Age First scheduling strategy: In this scheduling strategy, the source with the maximum age is served the first among all other sources. Ties may be broken arbitrarily.

We use π to represent a scheduling strategy. For simplicity, let π_{MAF} represent the MAF scheduling strategy. Moreover, let $f \triangleq (S_1, S_2, \ldots)$ denote a sampling strategy (i.e., the packet generation times). Hence, we have $d = (\pi, f)$, i.e., $d = (\pi, f)$ implies that a decision policy d follows the scheduling strategy π and the sampling strategy f. The age performance resulting from following π_{MAF} strategy is characterized as follows.

THEOREM 3.2. *Fix the sampling strategy f in* (4) *and* (5)*, then the MAF scheduling strategy minimizes the TaPA and the TaA among all possible scheduling strategies, i.e., for all f and π, we have*

$$\Delta_{peak}(\pi_{MAF}, f) \le \Delta_{peak}(\pi, f), \tag{8}$$

and

$$\Delta_{avg}(\pi_{MAF}, f) \le \Delta_{avg}(\pi, f). \tag{9}$$

PROOF. One of the key ideas of the proof is as follows. The sampling strategy is fixed. Thus, the packets generation times are fixed, and we only control from which source a packet is generated. We couple the policies such that the packet delivery times are fixed under all decision policies. Since we follow the MAF scheduling strategy, after each delivery, a source with maximum age becomes the source with minimum age among the m sources. Under any arbitrary scheduling strategy, a packet can be generated from any source, which is not necessary the one with maximum age, and the chosen source becomes the one with minimum age among the m sources after the delivery. Thus, following the MAF scheduling strategy provides a better age performance compared to following any other scheduling strategy. For more detail, see Appendix A. □

Remark 1. The result in Theorem 3.2 can be extended to general $\mathcal{Y}$ and $\mathcal{Z}$, i.e., $\mathcal{Y}$ and $\mathcal{Z}$ can be any infinite sets. In other words, Theorem 3.2 holds for any arbitrary distributed service times, including continuous service times. This is because the proof of Theorem 3.2 does not depend on the service time distribution.

Theorem 3.2 tells us that, for any values of the waiting times imposed by the decision policy, the TaPA and the TaA resulting from following the MAF scheduling strategy are lower than those resulting from following any other scheduling strategy π. This theorem helps us conclude the separation principle and design the optimal sampling strategy independently from the optimal scheduling strategy. In other words, Theorem 3.2 helps in shrinking the decision policy space by limiting the scheduling strategy to the MAF strategy; and it remains to search for the optimal sampling strategy. Recall that, we use f to denote a sampling strategy that a decision policy can follow, i.e., $f \triangleq (Z_0, Z_1, \ldots)$. Without a confusion, we will use the term "sampling policy" or "sampler" to denote the sampling strategy that a decision policy can follow. We use $\mathcal{F}$ to denote the set of causal sampling policies in which the decisions are made based on the history and current information of the system. As a result of Theorem 3.2, our optimization problems reduce to the following

$$\bar{r}_{\text{opt}} \triangleq \min_{f\in\mathcal{F}} \Delta_{\text{peak}}(\pi_{\text{MAF}}, f). \tag{10}$$

$$\bar{g}_{\text{opt}} \triangleq \min_{f\in\mathcal{F}} \Delta_{\text{avg}}(\pi_{\text{MAF}}, f). \tag{11}$$

Next, we seek the optimal samplers for problems (10) and (11).

3.2 Optimal Sampler for Problem (10)

By fixing our scheduling strategy to the MAF strategy, the evolution of the age processes of the sources is as follows. The sampler may impose a waiting time Z_i before generating packet $i+1$ at time $S_{i+1} = D_i + Z_i$ from the source with the maximum age at time $t = D_i$. Packet $i+1$ is delivered at time $D_{i+1} = S_{i+1} + Y_{i+1}$ and the age of the source with maximum age drops to the minimum age with the value of Y_{i+1}, while the age processes of other sources increase linearly with time without change. This operation is repeated with time and the age processes evolve accordingly. An example of age processes evolution is shown in Fig. 3. Next, we show that the zero-wait sampler minimize the TaPA.

THEOREM 3.3. *The optimal sampler for Problem* (10) *is the zero-wait sampler, i.e., $Z_i = 0$ for all i.*

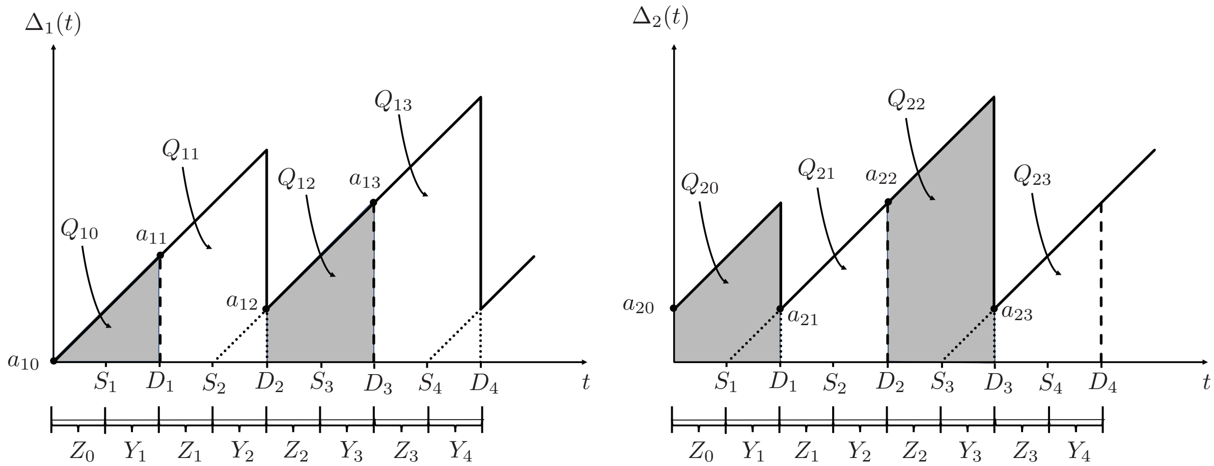
PROOF. We prove Theorem 3.3 by showing that the TaPA is an increasing function of the packets waiting times Z_i's. For notation simplicity, let a_{li} denote the age value for the source l at time D_i, i.e., $a_{li} = \Delta_l(D_i)$. Since the age process increases linearly with time when there is no packet delivery, we have

$$\Delta_{r_i}(D_i^-) = a_{r_i(i-1)} + Z_{i-1} + Y_i, \tag{12}$$

where $a_{r_i(i-1)} = \Delta_{r_i}(D_{i-1})$. Substituting by this in (4), we get

$$\Delta_{\text{peak}}(\pi_{\text{MAF}}, f) = \limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=1}^{n+1} a_{r_i(i-1)} + Z_{i-1} + Y_i\right]. \tag{13}$$

Since we follow the MAF scheduling strategy, the last serves of the source r_i before time D_{i-1} occurs at time D_{i-m}. Since the age



(a) The age evolution of source 1.

(b) The age evolution of source 2.

Figure 3: The age processes evolution in two sources information update system when the scheduling strategy is fixed to the MAF strategy. source 2 has a higher initial age than source 1, so source 2 starts service and packet 1 is generated from source 2, which is delivered at time D_1. Then, source 1 is served and packet 2 is generated from source 1, which is delivered at time D_2. The operation is repeated with time and the age processes evolve accordingly.

process increases linearly if there is no packet delivery, we have

$$a_{r_i(i-1)} = D_{i-1} - D_{i-m} + Y_{i-m}, \tag{14}$$

where Y_{i-m} is the age value of the source r_i at time D_{i-m}, i.e., $\Delta_{r_i}(D_{i-m}) = Y_{i-m}$. Note that $D_{i-1} = Y_{i-1} + Z_{i-2} + D_{i-2}$. Repeating this, we can express $(D_{i-1} - D_{i-m})$ in terms of Z_i's and Y_i's, and hence we get

$$a_{r_i(i-1)} = \sum_{k=1}^{m} Y_{i-k} + \sum_{k=2}^{m} Z_{i-k}. \tag{15}$$

For example, in Fig. 3, we have $a_{22} = Y_1 + Z_1 + Y_2$. Note that the regularity of (15) occurs after the first m transmissions (i.e., (15) is valid after the first m transmissions). For simplicity, we can omit the first m peaks in (13) ($\sum_{i=1}^{m} \Delta_{r_i}(D_i^-)$), as $\lim_{n\to\infty} \frac{\sum_{i=1}^{m} \Delta_{r_i}(D_i^-)}{n} = 0$ (observe that $\mathcal{Z}$ and $\mathcal{Y}$ are bounded), and this will not affect the value of the TaPA. This with substituting by (15) in (13), we get

$$\Delta_{\text{peak}}(\pi_{\text{MAF}}, f) = \limsup_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\sum_{i=m+1}^{n+1}\left(\sum_{k=0}^{m} Y_{i-k} + \sum_{k=1}^{m} Z_{i-k}\right)\right]. \tag{16}$$

From (16), it follows that the TaPA is an increasing function of the waiting times. This completes the proof. □

Remark 2. Similar to Theorem 3.2, the result in Theorem 3.3 can be extended to general $\mathcal{Y}$ and $\mathcal{Z}$, i.e., $\mathcal{Y}$ and $\mathcal{Z}$ can be any infinite bounded sets.

By this we conclude that the optimal scheduling strategy and sampling strategy for problem (6) are the MAF scheduling strategy and zero-wait sampling strategy, respectively.

3.3 Optimal Sampler for Problem (11)

Although the zero-wait sampler is the optimal sampler for minimizing the TaPA, it is not clear whether it also minimizes the TaA. This is because the latter metric may not be a non-decreasing function of the waiting times as we will see later. Next, we derive the TaA when the MAF scheduling strategy is followed and provide and equivalent mapping for problem (11).

3.3.1 Equivalent Mapping of Problem (11). We start by deriving the TaA when the decision policy follows the MAF scheduling strategy. We decompose the area under each curve $\Delta_l(t)$ into a sum of disjoint geometric parts. Observing Fig. 3, this area in the time interval $[0, D_n]$, where $D_n = \sum_{i=0}^{n-1} Z_i + Y_{i+1}$, can be seen as the concatenation of the areas Q_{li}, $0 \le i \le n-1$. Thus, we have

$$\int_0^{D_n} \Delta_l(t)dt = \sum_{i=0}^{n-1} Q_{li}. \tag{17}$$

Recall that we use a_{li} to denote the age value for the source l at time D_i, i.e., $a_{li} = \Delta_l(D_i)$. Then, as seen in Fig. 3, Q_{li} can be expressed as

$$Q_{li} = a_{li}(Z_i + Y_{i+1}) + \frac{1}{2}(Z_i + Y_{i+1})^2. \tag{18}$$

Using this with (17), we get

$$\sum_{l=1}^{m}\int_0^{D_n} \Delta_l(t)dt = \sum_{i=0}^{n-1} A_i(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2, \tag{19}$$

where $A_i = \sum_{l=1}^{m} a_{li}$. The TaA can be written as

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[A_i(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right]}{\sum_{i=0}^{n-1} \mathbb{E}\left[Z_i + Y_{i+1}\right]}. \tag{20}$$

Using this, the optimal sampling problem for minimizing the TaA, given that the scheduling strategy is fixed to the MAF strategy, can be formulated as

$$\bar{g}_{\text{opt}} \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[A_i(Z_i+Y_{i+1})+\frac{m}{2}(Z_i+Y_{i+1})^2\right]}{\sum_{i=0}^{n-1}\mathbb{E}[Z_i+Y_{i+1}]}. \quad (21)$$

Since $\mathcal{Z}$ and $\mathcal{Y}$ are bounded, $\bar{g}_{\text{opt}}$ is bounded as well. Note that problem (21) is hard to solve in the current form. Therefore, we provide an equivalent mapping for it. We consider the following optimization problem with a parameter $\beta \geq 0$:

$$p(\beta) \triangleq \min_{f\in\mathcal{F}} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[(A_i-\beta)(Z_i+Y_{i+1})+\frac{m}{2}(Z_i+Y_{i+1})^2\right], \quad (22)$$

where $p(\beta)$ is the optimal value of (22).

Lemma 3.4. *The following assertions are true:*

(i). $\bar{g}_{opt} \gtreqqless \beta$ *if and only if* $p(\beta) \gtreqqless 0$.

(ii). *If* $p(\beta)=0$*, then the optimal sampling policies that solve* (21) *and* (22) *are identical.*

Proof. The proof of Lemma 3.4 is similar to the proof of [29, Lemma 2]. The difference is that the regenerative property of the inter-sampling times is used to prove the result in [29]; instead, we use the boundedness of the inter-sampling times to prove the result. For the sake of completeness, we modify the proof accordingly and provide it in Appendix B. □

As a result of Lemma 3.4, the solution to (21) can be obtained by solving (22) and seeking a β_{opt} such that $p(\beta_{\text{opt}})=0$. Lemma 3.4 helps us to formulate our optimization problem as a DP problem. Without Lemma 3.4, it will be hard to formulate (21) as DP problem or solve it optimally. Next, we use the DP technique to solve problem (22).

3.3.2 The DP problem of (22). Following the methodology proposed in [7], when $\beta=\beta_{\text{opt}}$, the optimization problem (22) is equivalent to an average cost per stage DP problem. According to [7], we describe the components of our DP problem in detail below.

- **States:** At stage[2] i, the system state is specified by
$$\mathbf{s}(i) = (a_{[1]i},\ldots,a_{[m]i}). \quad (23)$$
We use $\mathcal{S}$ to denote the state-space including all possible states. Notice that $\mathcal{S}$ is finite and bounded because $\mathcal{Z}$ and $\mathcal{Y}$ are finite and bounded. Also, this implies that A_i's are uniformly bounded, i.e., there exists $\Lambda\in\mathbb{R}^+$ such that $A_i\leq\Lambda$ for all i.
- **Control action:** At stage i, the action that is taken by the sampler is $Z_i\in\mathcal{Z}$. Recall that $Z_i\leq M$ for all $i\geq 0$.
- **Random disturbance:** In our model, the random disturbance occurring at stage i is Y_{i+1}, which is independent of the system state and the control action.
- **Transition probabilities:** If the control $Z_i=z$ is applied at stage i and the service time of packet $i+1$ is $Y_{i+1}=y$, then the evolution of the system state from $\mathbf{s}(i)$ to $\mathbf{s}(i+1)$ is as follows.
$$\begin{aligned} &a_{[m]i+1}=y,\\ &a_{[l]i+1}=a_{[l+1]i}+z+y,\ l=1,\ldots,m-1. \end{aligned} \quad (24)$$
We let $\mathbb{P}_{\mathbf{ss}'}(z)$ denote the transition probabilities
$$\mathbb{P}_{\mathbf{ss}'}(z)=\mathbb{P}(\mathbf{s}(i+1)=\mathbf{s}'|\mathbf{s}(i)=\mathbf{s},Z_i=z),\ \mathbf{s},\mathbf{s}'\in\mathcal{S}.$$
When $\mathbf{s}=(a_{[1]},\ldots,a_{[m]})$ and $\mathbf{s}'=(a'_{[1]},\ldots,a'_{[m]})$, the law of the transition probability is given by
$$\mathbb{P}_{\mathbf{ss}'}(z)=\begin{cases}\mathbb{P}(Y_{i+1}=y) & \text{if } a'_{[m]}=y \text{ and}\\ & a'_{[l]}=a_{[l+1]}+z+y \text{ for } l\neq m;\\ 0 & \text{else.}\end{cases} \quad (25)$$
- **Cost Function:** Each time the system is in stage i and control Z_i is applied, we incur a cost
$$\begin{aligned} C(\mathbf{s}(i),Z_i=z,Y_{i+1}) =& (A_i-\beta_{\text{opt}})(z+Y_{i+1})+\\ &\frac{m}{2}(z^2+2zY_{i+1}+Y_{i+1}^2). \end{aligned} \quad (26)$$
To simplify notation, we use the expected cost $C(\mathbf{s}(i),Z_i)$ as the cost per stage, i.e.,
$$C(\mathbf{s}(i),Z_i)=\mathbb{E}_{Y_{i+1}}\left[C(\mathbf{s}(i),Z_i=z,Y_{i+1})\right], \quad (27)$$
where $\mathbb{E}_{Y_{i+1}}$ is the expectation with respect to Y_{i+1}. Hence, we have
$$\begin{aligned} C(\mathbf{s}(i),Z_i=z) =& (A_i-\beta_{\text{opt}})(z+\mathbb{E}[Y])+\\ &\frac{m}{2}(z^2+2z\mathbb{E}[Y]+\mathbb{E}\left[Y^2\right]), \end{aligned} \quad (28)$$
where we have used the fact that Z_i and Y_{i+1} are independent, and the random variable Y has the same distribution as the service times Y_i's. It is important to note that there exists $c\in\mathbb{R}^+$ such that $|C(\mathbf{s}(i),Z_i)|\leq c$ for all $\mathbf{s}(i)\in\mathcal{S}$ and $Z_i\in\mathcal{Z}$. This is because $\mathcal{Z}$, $\mathcal{Y}$, $\mathcal{S}$, and β_{opt} are bounded.

In general, the average cost per stage under a sampling policy $f\in\mathcal{F}$ is given by

$$\limsup_{n\to\infty}\frac{1}{n}\mathbb{E}\left[\sum_{i=0}^{n-1}C(\mathbf{s}(i),Z_i)\right]. \quad (29)$$

We say that a sampling policy $f\in\mathcal{F}$ is *average-optimal* if it minimizes the average cost per stage in (29). Our objective is to find the average-optimal sampling policy. A policy f is called history-dependent if the control Z_i depends on the entire past history, i.e., it depends on $\mathbf{s}(0),\ldots,\mathbf{s}(i)$ and $Z_0,\ldots,Z_{i-1}$. A policy is stationary if $Z_i=Z_j$ whenever $\mathbf{s}(i)=\mathbf{s}(j)$ for any i,j. In addition, a randomized policy assigns a probability distribution over the control set such that it chooses a control randomly according to this distribution, while a deterministic policy selects an action with certainty. According to [7], there may not exist a stationary deterministic policy that is average-optimal. In the next theorem, we show that there is actually a stationary deterministic policy that is average-optimal to our problem.

Theorem 3.5. *There exist a scalar* λ *and a function* h *that satisfy the following Bellman's equation*

$$\lambda+h(\mathbf{s})=\min_{z\in\mathcal{Z}}\left[C(\mathbf{s},z)+\sum_{\mathbf{s}'\in\mathcal{S}}\mathbb{P}_{\mathbf{ss}'}(z)h(\mathbf{s}')\right], \quad (30)$$

where λ *is the optimal average cost per stage that is independent of the initial state* $\mathbf{s}(0)$ *and satisfies*

$$\lambda=\lim_{\alpha\to 1}(1-\alpha)J_\alpha(\mathbf{s}),\forall\mathbf{s}\in\mathcal{S}, \quad (31)$$

[2]Throughout this paper, we assume that stage i starts at time D_i and ends at time D_{i+1}.

and $h(\mathbf{s})$ is the relative cost function that, for any state $\mathbf{o}$, satisfies

$$h(\mathbf{s}) = \lim_{\alpha \to 1} (J_\alpha(\mathbf{s}) - J_\alpha(\mathbf{o})), \forall \mathbf{s} \in \mathcal{S}, \tag{32}$$

where $J_\alpha(\mathbf{s})$ is the optimal total expected α-discounted cost function, which is defined by

$$J_\alpha(\mathbf{s}) = \min_{f \in \mathcal{F}} \limsup_{n \to \infty} \mathbb{E}\left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i)\right], \mathbf{s}(0) = \mathbf{s} \in \mathcal{S}, \tag{33}$$

where $0 < \alpha < 1$ is the discount factor. Furthermore, there exists a stationary deterministic policy that attains the minimum in (30) for each $\mathbf{s} \in \mathcal{S}$ and is average-optimal.

Proof. See Appendix C. □

We can deduce from Theorem 3.5 that the optimal waiting time is a fixed function of the state $\mathbf{s}$. Next, we provide a useful property of the optimal sampling policy that can reduce the complexity of the optimal sampler algorithm.

3.3.3 Optimal Sampler Structure. The relative value iteration (RVI) algorithm [25, Section 9.5.3], [18, Page 171] can be used to solve the Bellman's equation (30). Starting with an arbitrary state $\mathbf{o}$, a single iteration for the RVI algorithm is given as follows:

$$\begin{aligned} Q_{n+1}(\mathbf{s}, z) &= C(\mathbf{s}, z) + \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{s}\mathbf{s}'}(z) h_n(\mathbf{s}'), \\ J_{n+1}(\mathbf{s}) &= \min_{z \in \mathcal{Z}} (Q_{n+1}(\mathbf{s}, z)), \\ h_{n+1}(\mathbf{s}) &= J_{n+1}(\mathbf{s}) - J_{n+1}(\mathbf{o}), \end{aligned} \tag{34}$$

where $Q_{n+1}(\mathbf{s}, z)$, $J_n(\mathbf{s})$, and $h_n(\mathbf{s})$ denote the state action value function, value function, and relative value function for iteration n, respectively. At the beginning, we set $J_0(\mathbf{s}) = 0$ for all $\mathbf{s} \in \mathcal{S}$, and then we repeat the iteration of the RVI algorithm as described before.

The complexity of the RVI algorithm is high due to many sources (i.e., curse of dimensionatlity [23]). Thus, we need to simplify the RVI algorithm. To that end, we show that the optimal sampler has a threshold property that can reduce the complexity of the RVI algorithm.

Theorem 3.6. *Let A_s be the sum of the age values of state $\mathbf{s}$. Then, the optimal waiting time of any state $\mathbf{s}$ with $A_s \geq (\beta_{opt} - m\mathbb{E}[Y])$ is zero.*

Proof. See Appendix D. □

We can exploit Theorem 3.6 to reduce the complexity of the RVI algorithm as follows. The optimal waiting time for any state $\mathbf{s}$ whose $A_s \geq (\beta_{\text{opt}} - m\mathbb{E}[Y])$ is zero. Thus, we need to solve (34) only for the states whose $A_s \leq (\beta_{\text{opt}} - m\mathbb{E}[Y])$. As a result, we reduce the number of computations required along the system state space, which reduce the complexity of the RVI algorithm. Note that β_{opt} can be obtained using the bisection method or any other one-dimensional search methods. Combining this with the result of Theorem 3.6 and the RVI algorithm, we propose Algorithm 1, where z_s^* is the optimal waiting time for state $\mathbf{s}$. Note that the value of u in Algorithm 1 can be initialized to the value of the TaA of the zero-wait sampler (as the TaA of the zero-wait sampler provides an upper bound on the optimal TaA), which can be easily calculated. The RVI algorithm and Whittle's methodology have been used in literature to obtain the optimal age scheduler in a time-slotted multi-source networks (e.g.,[13, 14]). However, their proposed algorithms cannot handle random transmission times (which could be more than one time slot). Our proposed scheduling and sampling strategies can minimize the ToA in multi-source network with any random transmission time (with a discrete distribution).

Although the work in [31] provided the exact solution of the optimal sampling problem for minimizing the age in single source systems, its results hold only when there is a bounded on the waiting times. In this paper, we show that we can indeed generalize our results and eliminate the upper bound on the waiting times, M, as follows.

Lemma 3.7. *Let $\bar{g}_{opt}^{\infty}$ and $\bar{g}_{opt}^{M}$ represent the optimal TaA when the upper bound on the waiting times is ∞ and M, respectively. Also, let $\mathcal{Z}^\infty$ and $\mathcal{Z}^M$ represent the control space when the upper bound on the waiting times is ∞ and M, respectively. If $\mathcal{Z}^M \subseteq \mathcal{Z}^\infty$ for all values of M, then we have*

$$\bar{g}_{opt}^{\infty} = \lim_{M \to \infty} \bar{g}_{opt}^{M}. \tag{35}$$

Proof. See Appendix E. □

In a part of the proof of Lemma 3.7, we show that there is a limit after which increasing the bound M does not affect the optimal sampler. Hence, according to Lemma 3.7 and this part of the proof, we can get the value of $\bar{g}_{\text{opt}}^{\infty}$ using Algorithm 1 as follows. Starting from a large enough value of the bound M, we increase it gradually and run Algorithm 1 for each value. Once there is no change in the resulted solution (the optimal sampler) from Algorithm 1, then this is the solution that can obtain $\bar{g}_{\text{opt}}^{\infty}$. This solution is the optimal one among all values of M.

4 BELLMAN'S EQUATION APPROXIMATION

In this section, we provide an approximate analysis for Bellman equation in (30) in order to find a simple algorithm to solve problem (22). For a given state $\mathbf{s}$, we denote the next state given z and y by $\mathbf{s}'(z, y)$. From the state evolution in (24) and the transition probability equation (25), Bellman's equation in (30) can be rewritten as

$$\lambda = \min_z \left[C(\mathbf{s}, z) + \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y)(h(\mathbf{s}'(z, y)) - h(\mathbf{s}))\right]. \tag{36}$$

Although $h(\mathbf{s})$ is discrete, we could interpolate the value of $h(\mathbf{s})$ between the discrete values so that it is differentiable by following the same approach in [8] and [36]. Let $\mathbf{s} = (a_{[1]}, \ldots, a_{[m]})$, then using the first order Taylor approximation around a state $\mathbf{t} = (a_{[1]}^t, \ldots, a_{[m]}^t)$ (some fixed state), we get

$$h(\mathbf{s}) \approx h(\mathbf{t}) + \sum_{l=1}^{m} (a_{[l]} - a_{[l]}^t) \frac{\partial h(\mathbf{t})}{\partial a_{[l]}}. \tag{37}$$

Again, we use the first order Taylor approximation around the state $\mathbf{t}$, together with the state evolution in (24), to get

$$h(\mathbf{s}'(z, y)) \approx h(\mathbf{t}) + (y - a_{[m]}^t) \frac{\partial h(\mathbf{t})}{\partial a_{[m]}} + \sum_{l=1}^{m-1} (a_{[l+1]} - a_{[l]}^t + z + y) \frac{\partial h(\mathbf{t})}{\partial a_{[l]}}. \tag{38}$$

Algorithm 1: Optimal sampler algorithm.

given $l = 0$, sufficiently large u, tolerance $\epsilon_1 > 0$, tolerance $\epsilon_2 > 0$;
while $u - l > \epsilon_1$ **do**
  $\beta_{opt} = \frac{l+u}{2}$;
  $J(\mathbf{s}) = 0$, $h(\mathbf{s}) = 0$, $h_{\text{old}}(\mathbf{s}) = 0$ for all states $\mathbf{s} \in \mathcal{S}$;
  while $\max_{\mathbf{s}\in\mathcal{S}} |h(\mathbf{s}) - h_{old}(\mathbf{s})| > \epsilon_2$ **do**
    for *each* $\mathbf{s} \in \mathcal{S}$ **do**
      if $A_s \geq (\beta_{opt} - m\mathbb{E}[Y])$ **then**
        $z_s^* = 0$;
      else
        $z_s^* = \text{argmin}_{z\in\mathcal{Z}} C(\mathbf{s}, z) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z)h(\mathbf{s}')$;
      end
      $J(\mathbf{s}) = C(\mathbf{s}, z_s^*) + \sum_{\mathbf{s}'\in\mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z_s^*)h(\mathbf{s}')$;
    end
    $h_{\text{old}}(\mathbf{s}) = h(\mathbf{s})$;
    $h(\mathbf{s}) = J(\mathbf{s}) - J(\mathbf{o})$;
  end
  if $J(\mathbf{o}) \geq 0$ **then**
    $u = \beta_{\text{opt}}$;
  else
    $l = \beta_{\text{opt}}$;
  end
end

From (37) and (38), we get

$$h(\mathbf{s}'(z,y)) - h(\mathbf{s}) \approx (y - a_{[m]})\frac{\partial h(\mathbf{t})}{\partial a_{[m]}} + \sum_{l=1}^{m-1}(a_{[l+1]} - a_{[l]} + z + y)\frac{\partial h(\mathbf{t})}{\partial a_{[l]}}.$$

This implies that

$$\sum_{y\in\mathcal{Y}} \mathbb{P}(Y=y)(h(\mathbf{s}'(z,y)) - h(\mathbf{s})) \approx (\mathbb{E}[Y] - a_{[m]})\frac{\partial h(\mathbf{t})}{\partial a_{[m]}} + \sum_{l=1}^{m-1}(a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y])\frac{\partial h(\mathbf{t})}{\partial a_{[l]}}. \quad (39)$$

Using (36) with (39), we can get the following approximated Bellman's equation.

$$\lambda \approx \min_z (A_s - \beta_{\text{opt}})(z + \mathbb{E}[Y]) + \frac{m}{2}((z)^2 + 2z\mathbb{E}[Y] + \mathbb{E}\left[Y^2\right]) + (\mathbb{E}[Y] - a_{[m]})\frac{\partial h(\mathbf{t})}{\partial a_{[m]}} + \sum_{l=1}^{m-1}(a_{[l+1]} - a_{[l]} + z + \mathbb{E}[Y])\frac{\partial h(\mathbf{t})}{\partial a_{[l]}},$$

where A_s is the sum of the age values of state $\mathbf{s}$. A necessary condition for minimizing the RHS of the previous equation is to set its derivative to zero. We get

$$A_s - \beta_{\text{opt}} + mz + m\mathbb{E}[Y] + \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{t})}{\partial a_{[l]}} = 0. \quad (40)$$

Rearranging (40), we get

$$\hat{z}_s^* = \left[\frac{\beta_{\text{opt}} - m\mathbb{E}[Y] - \sum_{l=1}^{m-1} \frac{\partial h(\mathbf{t})}{\partial a_{[l]}}}{m} - \frac{A_s}{m}\right]^+, \quad (41)$$

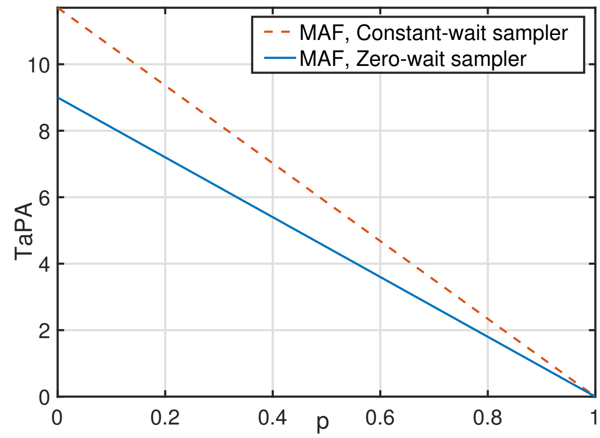


Figure 4: TaPA versus transmission probability p for an update system with $m = 3$ sources.

where $\hat{z}_s^*$ is the optimal solution of the approximated Bellman's equation for state $\mathbf{s}$. Note that the term $\sum_{i=1}^{m-1} \frac{\partial h(\mathbf{t})}{\partial a_{[i]}}$ is constant. Hence, (41) can be written as

$$\hat{z}_s^* = \left[th - \frac{A_s}{m}\right]^+, \quad (42)$$

where we have used Lemma 3.7 to eliminate the upper bound M in (42) (or simply M can be set to be greater than the optimal threshold in (42)). The solution in (42) is in the form of the water-filling solution as we compare a fixed threshold with the average age of a state $\mathbf{s}$. The solution in (42) suggests that the water-filling solution can approximate the optimal solution of the original Bellman's equation in (30). The optimal threshold in (42) can be obtained using a golden-section method [24]. We evaluate the performance of the water-filling solution obtained in (42) in the next section.

5 NUMERICAL RESULTS

We present some numerical results to verify our theoretical results. we consider an information update system with $m = 3$ sources. The packet transmission times are either 0 or 3 with probability p and $1-p$, respectively. We use "RAND" to represent the random scheduling strategy in which a source is chosen randomly to be served. Also, we use "constant-wait sampler" to represent the sampler that imposes a constant waiting time after each delivery, i.e., $Z_i = const$ for all i with $const = .3\mathbb{E}[Y]$. Finally, we use "optimal sampler" to refer to the sampler that results from Algorithm 1.

Fig. 4 illustrates the TaPA versus the transmission probability p. As we can observe, with fixing the scheduling strategy to the MAF strategy, the zero-wait sampler provides a lower TaPA compared to the constant-wait sampler. This observation agrees with Theorem 3.3. However, as we will see later, zero-wait sampler does not always minimize the TaA.

Fig. 5 illustrates the TaA versus the transmission probability p. For the zero-wait sampler, we find that following the MAF scheduling strategy provides a lower TaA than that is resulting from following the RAND scheduling strategy. This agrees with Theorem 3.2. Moreover, when the scheduling strategy is fixed to the

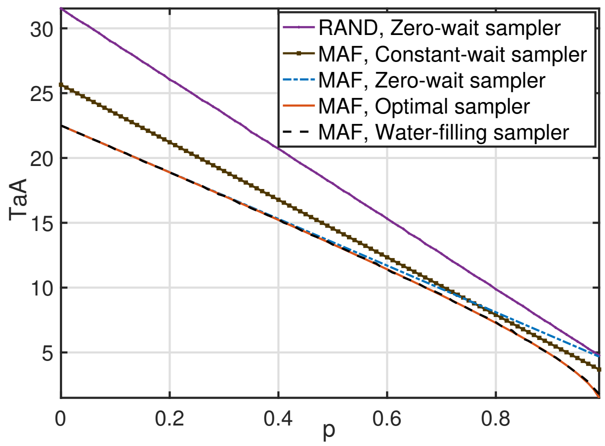


Figure 5: TaA versus transmission probability p for an update system with $m = 3$ sources.

MAF strategy, we find that the TaA results from the optimal sampler is lower than those result from the zero-wait sampler and the constant-wait sampler. This observation implies the following: i) The zero-wait sampler does not necessarily minimize the TaA, ii) optimizing the scheduling strategy only is not enough to minimize the TaA and we have to optimize both the scheduling strategy and the sampling strategy together to minimize the TaA. Finally, as we can observe that the TaA resulting from the water-filling sampler almost coincides on the TaA resulting from the optimal sampler.

6 CONCLUSIONS

In this paper, we studied the problem of finding the optimal decision policy that controls the packet generation times and schedules source transmissions to minimize the TaPA and TaA in multi-source information update system. We showed that our optimization problem has an important separation principle: the optimal scheduling strategy and the optimal sampling strategy can be designed independently. In particular, using the stochastic ordering technique, we proved that following the MAF scheduling strategy can provide a better age performance compared to following any other scheduling strategy, given that the sampling strategy is fixed. Later, we sought the optimal sampler for minimizing TaPA and TaA, given that the decision policy follows the MAF scheduling strategy. Although the zero-wait sampler was shown to be optimal for minimizing the TaPA, it did not always minimize the TaA. We used the DP technique to obtain the optimal sampler for minimizing the TaA and show that there exists a stationary deterministic sampler that can achieve optimality. Then, we showed that the optimal sampler has a threshold property that can reduce the complexity of the RVI algorithm. Finally, we provided an approximate analysis of Bellman's equation and showed that the water-filling solution can approximate the optimal sampler for minimizing the TaA. The numerical result showed that the performance of the water-filling solution is almost the same as that of the optimal sampler.

ACKNOWLEDGMENTS

This work has been supported in part by NSF grant CCF-1813050 and ONR grant N00014-17-1-2417.

REFERENCES

[1] B. Adelberg, H. Garcia-Molina, and B. Kao. 1995. Applying update streams in a soft real-time database system. In *ACM SIGMOD Record*, Vol. 24. 245–256.
[2] T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu. 2015. Age of Information under Energy Replenishment Constraints. In *Proc. Info. Theory and Appl. Workshop*.
[3] A. M. Bedewy, Y. Sun, and N. B. Shroff. 2016. Optimizing data freshness, throughput, and delay in multi-server information-update systems. In *Proc. IEEE ISIT*. 2569–2573.
[4] A. M. Bedewy, Y. Sun, and N. B. Shroff. 2017. Age-optimal information updates in multihop networks. In *Proc. IEEE ISIT*. 576–580.
[5] A. M. Bedewy, Y. Sun, and N. B. Shroff. submitted to *IEEE Trans. Inf. Theory*, 2017. Minimizing the Age of the Information through Queues. (submitted to *IEEE Trans. Inf. Theory*, 2017). https://arxiv.org/abs/1709.04956.
[6] A. M. Bedewy, Y. Sun, and N. B. Shroff. submitted to *IEEE/ACM Transactions on Networking*, 2018. The Age of Information in Multihop Networks. (submitted to *IEEE/ACM Transactions on Networking*, 2018). arXiv preprint arXiv:1712.10061.
[7] D. P. Bertsekas. 2001. *Dynamic Programming and Optimal Control, 2nd ed.* Vol. 2. Athena Scientific, Belmont, MA.
[8] I. Bettesh and S. Shamai. *IEEE Trans. Inf. Theory*, 2006. Optimal power and rate control for minimal average delay: The single-user case. 52, 9 (*IEEE Trans. Inf. Theory*, 2006), 4115–4141.
[9] J. Cho and H. Garcia-Molina. 2000. Synchronizing a database to improve freshness. In *ACM SIGMOD Record*, Vol. 29. 117–128.
[10] M. Costa, M. Codreanu, and A. Ephremides. 2016. On the Age of Information in Status Update Systems With Packet Management. IEEE Trans. Inf. Theory 62, 4 (2016), 1897–1910.
[11] L. Golab, T. Johnson, and V. Shkapenyuk. 2009. Scheduling Updates in a Real-Time Stream Warehouse. In *Proc. IEEE 25th International Conference on Data Engineering*. 1207–1210.
[12] Q. He, D. Yuan, and A. Ephremides. 2017. Optimal link scheduling for age minimization in wireless systems. IEEE Trans. Inf. Theory (2017).
[13] Y. Hsu. 2018. Age of information: Whittle index for scheduling stochastic arrivals. *arXiv preprint arXiv:1801.03422* (2018).
[14] Y. Hsu, E. Modiano, and L. Duan. 2017. Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals. *arXiv preprint arXiv:1712.07419* (2017).
[15] L. Huang and E. Modiano. 2015. Optimizing age-of-information in a multi-class queueing system. In *Proc. IEEE ISIT*. 1681–1685.
[16] I. Kadota, A. Sinha, and E. Modiano. 2018. Optimizing age of information in wireless networks with throughput constraints. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 1844–1852.
[17] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano. 2016. Minimizing the age of information in broadcast wireless networks. In *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*. IEEE, 844–851.
[18] L. P. Kaelbling and L. P. Kaelbling. 1996. *Recent advances in reinforcement learning*. Springer.
[19] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides. 2016. Effect of Message Transmission Path Diversity on Status Age. *IEEE Trans. Inf. Theory* 62, 3 (2016), 1360–1374.
[20] S. Kaul, R. D. Yates, and M. Gruteser. 2012. Real-time status: How often should one update?. In *Proc. IEEE INFOCOM*. 2731–2735.
[21] S. Kaul, R. D. Yates, and M. Gruteser. 2012. Status Updates Through Queues. In *Conf. on Info. Sciences and Systems*.
[22] N. Pappas, J. Gunnarsson, L. Kratz, M. Kountouris, and V. Angelakis. 2015. Age of information of multiple sources with queue management. In *Proc. IEEE ICC*. 5935–5940.
[23] W. B. Powell. 2007. *Approximate Dynamic Programming: Solving the curses of dimensionality*. Vol. 703. John Wiley & Sons.
[24] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. Golden section search in one dimension. *Numerical Recipes in C: The Art of Scientific Computing* (1992), 2.
[25] M. L. Puterman. 2005. Markov Decision Processes: Discrete Stochastic Dynamic Programming (Wiley Series in Probability and Statistics). (2005).
[26] Y. Sang, B. Li, and B. Ji. 2017. The Power of Waiting for More than One Response in Minimizing the Age-of-Information. (2017). https://arxiv.org/abs/1704.04848.
[27] L. I. Sennott. 1989. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research* 37, 4 (1989), 626–633.
[28] M. Shaked and J. G. Shanthikumar. 2007. *Stochastic orders*. Springer Science & Business Media.

[29] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu. 2017. Remote estimation of the Wiener process over a channel with random delay. In *Proc. IEEE ISIT*. 321–325.

[30] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella. 2018. Age-optimal updates of multiple information flows. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 136–141.

[31] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff. *IEEE Trans. Inf. Theory*, 2017. Update or Wait: How to Keep Your Data Fresh. (*IEEE Trans. Inf. Theory*, 2017).

[32] R. Talak, I. Kadota, S. Karaman, and E. Modiano. 2018. Scheduling policies for age minimization in wireless networks with unknown channel state. *arXiv preprint arXiv:1805.06752* (2018).

[33] R. Talak, S. Karaman, and E. Modiano. 2018. Distributed scheduling algorithms for optimizing information freshness in wireless networks. *arXiv preprint arXiv:1803.06469* (2018).

[34] R. Talak, S. Karaman, and E. Modiano. 2018. Optimizing age of information in wireless networks with perfect channel state information. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2018 16th International Symposium on*. IEEE, 1–8.

[35] R. Talak, S. Karaman, and E. Modiano. 2018. Optimizing information freshness in wireless networks under general interference constraints. *arXiv preprint arXiv:1803.06467* (2018).

[36] R. Wang and V. K. Lau. *IEEE Trans. Inf. Theory*, 2013. Delay-optimal two-hop cooperative relay communications via approximate mdp and distributive stochastic learning. (*IEEE Trans. Inf. Theory*, 2013).

[37] R. D. Yates. 2015. Lazy is Timely: Status Updates by an Energy Harvesting Source. In *Proc. IEEE Int. Symp. Inform. Theory*.

[38] R. D. Yates and S. Kaul. 2012. Real-time status updating: Multiple sources. In *Proc. IEEE ISIT*. 2666–2670.

[39] R. D. Yates and S. K. Kaul. 2017. Status updates over unreliable multiaccess channels. In *Proc. IEEE ISIT*. IEEE, 331–335.

A PROOF OF THEOREM 3.2

Let the vector $\mathbf{\Delta}_\pi(t) = (\Delta_{[1],\pi}(t), \ldots, \Delta_{[m],\pi}(t))$ denote the system state at time t when the scheduling strategy π is followed and $\{\mathbf{\Delta}_\pi(t), t \geq 0\}$ denote the state process when the scheduling strategy π is followed (for simplicity, we will call $\{\mathbf{\Delta}_\pi(t), t \geq 0\}$ the state process of the scheduling strategy π hereafter without confusion). For notation simplicity, let P represent the MAF scheduling strategy. Throughout the proof, we assume that $\mathbf{\Delta}_\pi(0^-) = \mathbf{\Delta}_P(0^-)$ for all π and the sampling strategy is fixed. The key step in the proof of Theorem 3.2 is the following lemma, where we compare the scheduling strategy P with any arbitrary scheduling strategy π.

Lemma A.1. *Suppose that $\mathbf{\Delta}_\pi(0^-) = \mathbf{\Delta}_P(0^-)$ for all scheduling strategy π and the sampling strategy is fixed, then we have*

$$\{\mathbf{\Delta}_P(t), t \geq 0\} \leq_{st} \{\mathbf{\Delta}_\pi(t), t \geq 0\} \tag{43}$$

We use a coupling and forward induction to prove Lemma A.1. For any scheduling strategy π, suppose that the stochastic processes $\widetilde{\mathbf{\Delta}}_P(t)$ and $\widetilde{\mathbf{\Delta}}_\pi(t)$ have the same stochastic laws as $\mathbf{\Delta}_P(t)$ and $\mathbf{\Delta}_\pi(t)$. The state processes $\widetilde{\mathbf{\Delta}}_P(t)$ and $\widetilde{\mathbf{\Delta}}_\pi(t)$ are coupled such that the packet service times are equal under both scheduling policies, i.e., Y_i's are the same under both scheduling policies. Such a coupling is valid since the service time distribution is fixed under all policies. Since the sampling strategy is fixed, such a coupling implies that the packet generation and delivery times are the same under both scheduling strategies. According to Theorem 6.B.30 of [28], if we can show

$$\mathbb{P}\left[\widetilde{\mathbf{\Delta}}_P(t) \leq \widetilde{\mathbf{\Delta}}_\pi(t), t \geq 0\right] = 1, \tag{44}$$

then (43) is proven. To ease the notational burden, we will omit the tildes on the coupled versions in this proof and just use $\mathbf{\Delta}_P(t)$ and $\mathbf{\Delta}_\pi(t)$. Next, we compare strategy P and strategy π on a sample path and prove (43) using the following lemma:

Lemma A.2 (Inductive Comparison). *Suppose that a packet with generation time S is delivered under the scheduling strategy P and the scheduling strategy π at the same time t. The system state of the scheduling strategy P is $\mathbf{\Delta}_P$ before the packet delivery, which becomes $\mathbf{\Delta}'_P$ after the packet delivery. The system state of the scheduling strategy π is $\mathbf{\Delta}_\pi$ before the packet delivery, which becomes $\mathbf{\Delta}'_\pi$ after the packet delivery. If*

$$\Delta_{[i],P} \leq \Delta_{[i],\pi}, i = 1, \ldots, m, \tag{45}$$

then

$$\Delta'_{[i],P} \leq \Delta'_{[i],\pi}, i = 1, \ldots, m. \tag{46}$$

Proof. Since only one source can be scheduled at a time and the scheduling strategy P is the MAF scheduling strategy, the packet with generation time S must be generated from the source with maximum age $\Delta_{[1],P}$, call it source l^*. In other words, the age of source l^* is reduced from the maximum age $\Delta_{[1],P}$ to the minimum age $\Delta'_{[m],P} = t - S$, and the age of the other $(m-1)$ sources remain unchanged. Hence,

$$\begin{aligned} &\Delta'_{[i],P} = \Delta_{[i+1],P}, i = 1, \ldots, m-1, \\ &\Delta'_{[m],P} = t - S. \end{aligned} \tag{47}$$

In the scheduling strategy π, this packet can be generated from any source. Thus, for all cases of strategy π, it must hold that

$$\Delta'_{[i],\pi} \geq \Delta_{[i+1],\pi}, i = 1, \ldots, m-1. \tag{48}$$

By combining (45), (47), and (48), we have

$$\Delta'_{[i],\pi} \geq \Delta_{[i+1],\pi} \geq \Delta_{[i+1],P} = \Delta'_{[i],P}, i = 1, \ldots, m-1. \tag{49}$$

In addition, since the same packet is also delivered under the scheduling strategy π, the source from which this packet is generated under policy π will have the minimum age after the delivery, i.e., we have

$$\Delta'_{[m],\pi} = t - S = \Delta'_{[m],P}. \tag{50}$$

By this, (46) is proven. □

Proof of Lemma A.1. Using the coupling between the system state processes, and for any given sample path of the packet service times, we consider two cases:

Case 1: When there is no packet delivery, the age of each source grows linearly with a slope 1.

Case 2: When a packet is delivered, the ages of the sources evolve according to Lemma A.2.

By induction over time, we obtain

$$\Delta_{[i],P}(t) \leq \Delta_{[i],\pi}(t), i = 1, \ldots, m, t \geq 0. \tag{51}$$

Hence, (44) follows which implies (43) by Theorem 6.B.30 of [28]. This completes the proof. □

Proof of Theorem 3.2. Since the TaPA and TaA for any scheduling policy π are the expectation of non-decreasing functionals of the process $\{\mathbf{\Delta}_\pi(t), t \geq 0\}$, (43) implies (8) and (9) using the properties of stochastic ordering [28]. This completes the proof. □

B PROOF OF LEMMA 3.4

Part (i) is proven in two steps:

Step 1: We will prove that $\bar{g}_{\text{opt}} \leq \beta$ if and only if $p(\beta) \leq 0$.

If $\bar{g}_{\text{opt}} \leq \beta$, there exists a sampling policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ that is feasible for (21) and (22), which satisfies

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}\left[A_i(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right]}{\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \leq \beta. \quad (52)$$

Hence,

$$\limsup_{n\to\infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[(A_i - \beta)(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right]}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]} \leq 0. \quad (53)$$

Since Z_i's and Y_i's are bounded and positive, we have ($\liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]$) is bounded and positive as well. By this, we get

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[(A_i - \beta)(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right] \leq 0. \quad (54)$$

Therefore, $p(\beta) \leq 0$.

In the revers direction, if $p(\beta) \leq 0$, then there exists a sampling policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ that is feasible for (21) and (22), which satisfies (54). This implies

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[A_i(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right]\\ &\leq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \beta\mathbb{E}[Z_i + Y_{i+1}],\end{aligned} \quad (55)$$

where we have used that $\limsup_{n\to\infty} \frac{-1}{n}\sum_{i=0}^{n-1} \beta\mathbb{E}[Z_i + Y_{i+1}] = -\liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \beta\mathbb{E}[Z_i+Y_{i+1}]$. By divide (55) by $\liminf_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{E}[Z_i + Y_{i+1}]$, we get (52). Hence, $\bar{g}_{\text{opt}} \leq \beta$. By this, we have proven that $\bar{g}_{\text{opt}} \leq \beta$ if and only if $p(\beta) \leq 0$.

Step 2: We need to prove that $\bar{g}_{\text{opt}} < \beta$ if and only if $p(\beta) < 0$. This statement can be proven by using the arguments in Step 1, in which "≤" should be replaced by "<". Finally, from the statement of Step 1, it immediately follows that $\bar{g}_{\text{opt}} > \beta$ if and only if $p(\beta) > 0$. This completes part (i).

Part(ii): We first show that each optimal solution to (21) is an optimal solution to (22). By the claim of part (i), $p(\beta) = 0$ is equivalent to $\bar{g}_{\text{opt}} = \beta$. Suppose that policy $f = (Z_0, Z_1, \ldots) \in \mathcal{F}$ is an optimal solution to (21). Then, $\Delta_{\text{avg}(\pi_{\text{MAF}}, f)} = \bar{g}_{\text{opt}} = \beta$. Applying this in the arguments of (52)-(54), we can show that policy f satisfies

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[(A_i - \beta)(Z_i + Y_{i+1}) + \frac{m}{2}(Z_i + Y_{i+1})^2\right] = 0. \quad (56)$$

This and $p(\beta) = 0$ imply that policy f is an optimal solution to (22).

Similarly, we can prove that each optimal solution to (22) is an optimal solution to (21). By this, part (ii) is proven.

C PROOF OF THEOREM 3.5

According to [7, Proposition 4.2.1 and Proposition 4.2.6], it is enough to show that for every two states $\mathbf{s}$ and $\mathbf{s}'$, there exists a stationary deterministic policy f such that for some k, we have

$$\mathbb{P}\left[\mathbf{s}(k) = \mathbf{s}'|\mathbf{s}(0) = \mathbf{s}, f\right] > 0. \quad (57)$$

From the state evolution equation (24), we can observe that any state in $\mathcal{S}$ can be represented in terms of the packet waiting and service times. This implies (57). To make it clearer, we consider the following case of the 3 sources system:

Assume that the elements of state $\mathbf{s}'$ are as follows

$$\begin{aligned}a'_{[1]} &= y_3 + z_2 + y_2 + z_1 + y_1,\\ a'_{[2]} &= y_3 + z_2 + y_2,\\ a'_{[3]} &= y_3,\end{aligned} \quad (58)$$

where y_i's and z_i's are any arbitrary elements in $\mathcal{Y}$ and $\mathcal{Z}$, respectively. Then, we will show that from any arbitrary state $\mathbf{s} = (a_{[1]}, a_{[2]}, a_{[3]})$, a sequence of service and waiting times can be followed to reach state $\mathbf{s}'$. If we have $Z_0 = z_1$, $Y_1 = y_1$, $Z_1 = z_1$, $Y_2 = y_2$, $Z_2 = z_2$, and $Y_3 = y_3$, then according to (24), we have in the first stage

$$\begin{aligned}a_{[1]1} &= a_{[2]} + z_1 + y_1,\\ a_{[2]1} &= a_{[3]} + z_1 + y_1,\\ a_{[3]1} &= y_1,\end{aligned} \quad (59)$$

and in the second stage, we have

$$\begin{aligned}a_{[1]2} &= a_{[3]} + z_1 + y_2 + z_1 + y_1,\\ a_{[2]2} &= y_2 + z_1 + y_1,\\ a_{[3]2} &= y_2,\end{aligned} \quad (60)$$

and in the third stage, we have

$$\begin{aligned}a_{[1]3} &= y_3 + z_2 + y_2 + z_1 + y_1 = a'_{[1]},\\ a_{[2]3} &= y_3 + z_2 + y_2 = a'_{[2]},\\ a_{[3]3} &= y_3 = a'_{[3]}.\end{aligned} \quad (61)$$

Hence, a stationary deterministic policy f can be designed to reach state $\mathbf{s}'$ from state $\mathbf{s}$ in 3 stages, if the aforementioned sequence of service times occurs. This implies that

$$\mathbb{P}\left[\mathbf{s}(3) = \mathbf{s}'|\mathbf{s}(0) = \mathbf{s}, f\right] = \prod_{i=1}^{3} \mathbb{P}(Y_i = y_i) > 0, \quad (62)$$

where we have used that Y_i's are *i.i.d.*[3] The previous argument can be generalized for any m sources system. In particular, a forward induction over m can be used to show the result, as (57) trivially holds for $m = 1$, and the previous argument can be used to show that (57) holds for any general m. This completes the proof.

D PROOF OF THEOREM 3.6

We prove Theorem 3.6 into two steps:

Step 1: We first address an infinite horizon discounted cost problem. Then, we connect it to the average cost per stage problem. In particular, we show that $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$, which together with (32) imply that $h(\mathbf{s})$ is non-decreasing in $\mathbf{s}$ as well.

Given an initial state $\mathbf{s}(0)$, the total expected discounted cost under a sampling policy $f \in \mathcal{F}$ is given by

$$J_\alpha(\mathbf{s}(0); f) = \limsup_{n\to\infty} \mathbb{E}\left[\sum_{i=0}^{n-1} \alpha^i C(\mathbf{s}(i), Z_i)\right], \quad (63)$$

[3] We assume that all elements in $\mathcal{Y}$ have a strictly positive probability, where the elements with zero probability can be removed without affecting the proof.

where $0 < \alpha < 1$ is the discount factor. The optimal total expected α-discounted cost function is defined by

$$J_\alpha(\mathbf{s}) = \min_{f \in \mathcal{F}} J_\alpha(\mathbf{s}; f), \ \mathbf{s} \in \mathcal{S}. \tag{64}$$

A policy is said to be α-optimal if it minimizes the total expected α-discounted cost. The discounted cost optimality equation of $J_\alpha(\mathbf{s})$ is discussed below.

Proposition D.1. *The optimal total expected α-discounted cost $J_\alpha(\mathbf{s})$ satisfies*

$$J_\alpha(\mathbf{s}) = \min_{z \in \mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_\alpha(\mathbf{s}'). \tag{65}$$

Moreover, a stationary deterministic policy that attains the minimum in equation (65) for each $\mathbf{s} \in \mathcal{S}$ will be an α-optimal policy. Also, let $J_{\alpha,0}(\mathbf{s}) = 0$ for all $\mathbf{s}$ and for any $n \geq 0$,

$$J_{\alpha,n+1}(\mathbf{s}) = \min_{z \in \mathcal{Z}} C(\mathbf{s}, z) + \alpha \sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}'). \tag{66}$$

Then, we have $J_{\alpha,n}(\mathbf{s}) \to J_\alpha(\mathbf{s})$ as $n \to \infty$ for every $\mathbf{s}$, and α.

Proof. Since we have bounded cost per stage, the proposition follows directly from [7, Proposition 1.2.2 and Proposition 1.2.3], and [27]. □

Next, we use the optimality equation (65) and the value iteration in (66) to prove that $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.

Lemma D.2. *The optimal total expected α-discounted cost function $J_\alpha(\mathbf{s})$ is non-decreasing in $\mathbf{s}$.*

Proof. We use induction on n in equation (66) to prove Lemma D.2. Obviously, the result holds for $J_{\alpha,0}(\mathbf{s})$.

Now, assume that $J_{\alpha,n}(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. We need to show that for any two states $\mathbf{s}_1$ and $\mathbf{s}_2$ with $\mathbf{s}_1 \leq \mathbf{s}_2$, we have $J_{\alpha,n+1}(\mathbf{s}_1) \leq J_{\alpha,n+1}(\mathbf{s}_2)$. First, we note that the expected cost per stage $C(\mathbf{s}, z)$ is non-decreasing in $\mathbf{s}$, i.e., we have

$$C(\mathbf{s}_1, z) \leq C(\mathbf{s}_2, z). \tag{67}$$

From the state evolution equation (24) and the transition probability equation (25), the second term of the right-hand side (RHS) of (66) can be rewritten as

$$\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) J_{\alpha,n}(\mathbf{s}') = \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'(z, y)), \tag{68}$$

where $\mathbf{s}'(z, y)$ is the next state from state $\mathbf{s}$ given the values of z and y. Also, according to the state evolution equation (24), if the next states of $\mathbf{s}_1$ and $\mathbf{s}_2$ for given values of z and y are $\mathbf{s}'_1(z, y)$ and $\mathbf{s}'_2(z, y)$, respectively, then we have $\mathbf{s}'_1(z, y) \leq \mathbf{s}'_2(z, y)$. This implies that

$$\sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'_1(z, y)) \leq \sum_{y \in \mathcal{Y}} \mathbb{P}(Y = y) J_{\alpha,n}(\mathbf{s}'_2(z, y)), \tag{69}$$

where we have used the induction assumption that $J_{\alpha,n}(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. Using (67), (69), and that fact that the minimum operator in (66) holds the non-decreasing property, we conclude that

$$J_{\alpha,n+1}(\mathbf{s}_1) \leq J_{\alpha,n+1}(\mathbf{s}_2). \tag{70}$$

This completes the proof. □

Step 2: We use Step 1 to prove Theorem 3.6. From Step 1, we have that $h(\mathbf{s})$ is non-decreasing in $\mathbf{s}$. Similar to Step 1, this implies that the second term of the right-hand side (RHS) of (30) ($\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) h(\mathbf{s}')$) is non-decreasing in $\mathbf{s}'$. Moreover, from the state evolution (24), we can notice that, for any state $\mathbf{s}$, the next state $\mathbf{s}'$ is increasing in z. This argument implies that the second term of the right-hand side (RHS) of (30) ($\sum_{\mathbf{s}' \in \mathcal{S}} \mathbb{P}_{\mathbf{ss}'}(z) h(\mathbf{s}')$) is increasing in z. Thus, the value of $z \in \mathcal{Z}$ that achieves the minimum value of this term is zero. If, for a given state $\mathbf{s}$, the value of $z \in \mathcal{Z}$ that achieves the minimum value of the cost function $C(\mathbf{s}, z)$ is zero, then $z = 0$ solves the RHS of (30). From (28), we observe that $C(\mathbf{s}, z)$ is convex in z. Also, the value of z that minimizes $C(\mathbf{s}, z)$ is $\frac{\beta_{\text{opt}} - A_s - m\mathbb{E}[Y]}{m}$. This implies that for any state $\mathbf{s}$ with $A_s \geq (\beta_{\text{opt}} - m\mathbb{E}[Y])$, $z = 0$ minimizes $C(\mathbf{s}, z)$ in the domain $\mathcal{Z}$. Hence, for any state $\mathbf{s}$ with $A_s \geq (\beta_{\text{opt}} - m\mathbb{E}[Y])$, $z = 0$ solves the RHS of (30). This completes the proof.

E PROOF OF LEMMA 3.7

We prove Lemma 3.7 into 3 steps.

Step 1: We show that when the upper bound on the waiting times keeps increasing, the optimal waiting times are still bounded. Let $z^*(\mathbf{s})$ be the optimal waiting time for the state $\mathbf{s}$. From Theorem 3.6, we know that when $A_s \geq \beta_{\text{opt}}$ (recall that $A_\mathbf{s}$ is the sum of ages of state $\mathbf{s}$), the optimal waiting time $z^*(\mathbf{s})$ is zero (this also can be deduced from problem (22)). Hence, we can restrict our focus on the age values whose summations are less than β_{opt} (which is a finite subset of the system state space). It is obvious that the optimal waiting time at any stage cannot increase without limit (i.e., goes to ∞) as the zero-wait sampler (the sampler that generates a packet once the previous one is delivered) can provide a lower TaA in this case. Thus, the optimal waiting times, when M is large enough, are upper bounded by the following bound.

$$B = \max\{z^*(\mathbf{s}) : A_\mathbf{s} < \beta_{\text{opt}}\}. \tag{71}$$

Observe that there is a limit after which increasing M just adds states with $A_s \geq \beta$, and hence does not affect the bound in (71).

Step 2: Let f^*_∞ and f^*_M be the optimal samplers when the upper bound on the waiting times is ∞ and M, respectively. Then, we show that $f^*_\infty = \lim_{M\to\infty} f^*_M$. From Step 1, we can conclude that there exists $n^* \geq B$ such that we have $f^*_M = f^*_N$ for any $M, N \geq n^*$. This implies that $f^*_\infty = \lim_{M\to\infty} f^*_M$. In other words, from Step 1 and problem (21), we can conclude that there is a limit (n^*) after which increasing the upper bound on the waiting times does not affect the optimal sampler.

Step 3: Finally, we show that $\bar{g}^\infty_{\text{opt}} = \lim_{M\to\infty} \bar{g}^M_{\text{opt}}$. From (21), we can write

$$\bar{g}^\infty_{\text{opt}} = \limsup_{n\to\infty} g^n(f^*_\infty), \tag{72}$$

where $g^n(f^*_\infty)$ is the total average age until the n-th delivery when the sampler follows the policy f^*_∞. Similarly, we can define

$$\bar{g}^M_{\text{opt}} = \limsup_{n\to\infty} g^n(f^*_M). \tag{73}$$

From Step 1 and 2, we can rewrite (72) as follows.

$$\bar{g}^\infty_{\text{opt}} = \limsup_{n\to\infty} \lim_{M\to\infty} g^n(f^*_M). \tag{74}$$

Notice that the TaA of the zero-wait sampler is bounded. This implies that there exists $G \in \mathbb{R}$ such that $|g^n(f_M^*)| \leq G$ for all M and n. Using bounded convergence theorem, we can interchange the limits in (74) to obtain

$$\bar{g}_{\text{opt}}^{\infty} = \lim_{M\to\infty} \limsup_{n\to\infty} g^n(f_M^*) = \lim_{M\to\infty} \bar{g}_{\text{opt}}^{M}, \tag{75}$$

which completes the proof.