

Cyclic Scheduler Design for Minimizing Age of Information in Massive Scale Networks Susceptible to Packet Errors

Sahan Liyanaarachchi¹, Graduate Student Member, IEEE, Sennur Ulukus², Fellow, IEEE, and Nail Akar³, Senior Member, IEEE

Abstract—In multi-source status update systems, sources need to be scheduled appropriately to maintain timely communication between each of the sources and the monitor. A *cyclic schedule* is an age-agnostic schedule in which the sources are served according to a fixed finite transmission pattern, which upon completion, repeats itself. Such a scheme has a low $O(1)$ runtime complexity, which is desirable in large networks. This paper's focus is on designing transmission patterns so as to be used in massive scale networking scenarios involving a very large number of sources, e.g., up to thousands of IoT sources, with service time requirements and weights being heterogeneous in nature. The goal is to minimize the weighted sum age of information (AoI), called weighted AoI, when transmitting users' packets over a channel susceptible to heterogeneous packet errors. The main tool we use is a stochastic modeling framework using either Markov chains (MC) or moment generating functions (MGF), by which we obtain the weighted AoI for a given transmission pattern, which is not straightforward in the presence of packet drops. Using this framework, we provide a lower bound on the weighted AoI for the particular case of two sources, and also an algorithm to attain this lower bound. Then, by using the same framework, we design a cyclic scheduler for general number of sources with reasonable complexity using convex optimization and well-established packet spreading algorithms, and comparatively evaluate the proposed algorithm and existing age-agnostic scheduling schemes for general number of sources (resp. two sources) when the lower bound is not available (resp. when it is available). We present extensive numerical results also involving comparisons with age-aware scheduling to validate the effectiveness of the proposed approach.

Index Terms—Age of information (AoI), cyclic schedulers, packet errors.

Received 20 June 2024; revised 3 December 2024; accepted 15 February 2025. Date of publication 5 March 2025; date of current version 21 May 2026. An earlier version of this paper was presented in part at the 2024 IEEE INFOCOM Workshop on Age and Semantics of Information and in part at the 58th Asilomar Conference on Signals, Systems and Computers. (Corresponding author: Sennur Ulukus.)

Sahan Liyanaarachchi and Sennur Ulukus are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: sahanl@umd.edu; ulukus@umd.edu).

Nail Akar is with the Department of Electrical and Electronics Engineering, Bilkent University, 06800 Ankara, Türkiye (e-mail: akar@ee.bilkent.edu.tr).

Communicated by E. Jorswieck, Associate Editor for Communications.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIT.2025.3548017>.

Digital Object Identifier 10.1109/TIT.2025.3548017

I. INTRODUCTION

NEED for timely status updates arises naturally in many remote monitoring and networked control systems [1], [2], [3]. A multi-source status update system consists of a monitor which receives status updates from multiple sources through a shared communications channel [4]. A common metric for quantifying timeliness in status update systems is the time average of the age of information (AoI) process, which we term as the mean AoI or average AoI [5], maintained at a certain device in the network, such as an intermediate node, a server, or the eventual destination, i.e., the monitor. In particular, the AoI process for source- n denoted by $\Delta_n(t)$, $t \geq 0$, is defined as $\Delta_n(t) = t - g_n(t)$, where $g_n(t)$ is the time of generation of the freshest update of source- n available at time t to the particular device of interest [6], [7], [8]. In this multi-source setting, status updates of some sources may be more critical than others, and therefore, weighted average of mean AoI values of the sources, termed weighted AoI, is often used as a system-wide metric for quantifying the timeliness of communication. To minimize the weighted AoI, source transmissions need to be scheduled appropriately. On the other hand, massive internet of things (IoT) networks are becoming commonplace that consist of a very large number of low cost and low power end users that report low frequency status updates to remote locations via 6G networks [9] or low power wide area networks (LPWAN) such as LoRaWAN [10]. Development of scalable schedulers to be deployed in massive scale networks for weighted AoI minimization is the main scope of this paper.

The first representative scenario (called Scenario A) for a massive IoT network we consider in this paper is an error-prone cellular wireless network with a server or a base station (BS) collecting time-sensitive information from a large number of users that are spread out at random locations in the wireless network. Each user is associated with an information source process, whose sample values are to be transmitted to the server which is in charge of forwarding these messages to a remote monitor, which may or may not be co-located with the server. Consequently, the terms “user” and “source” will be used interchangeably throughout the paper. In case of co-location, which is the scope of the current paper, AoI processes at both the server and the monitor are identical. However, these AoI processes may very well be different in the case of remotely located monitors, a situation which is left outside the scope of the current paper. Similar uplink scenarios

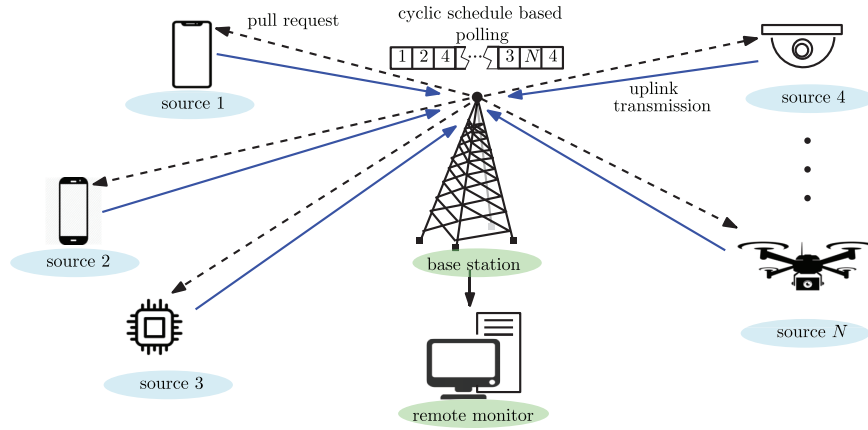


Fig. 1. Pull-based status update system (Scenario A) involving a base station (BS) sending polling messages to N users which in turn transmit their status update packets.

are investigated in [11], [12], and [13], where the status update messages are generated according to a random process, or periodically. In this paper, we consider the *generate-at-will* (GAW) model introduced in [14] which is a pull-based model where the server is to decide which source to collect the time-sensitive information from. Once the decision is made by the server, polling multiple access will be used by which the server sends a polling message (or query) to the scheduled user which subsequently samples the corresponding scheduled source process, generates a status update packet, and transmits to the server in the uplink direction, while avoiding collisions. Although random access is generally used as the channel access scheme for timely status updates [13], [15], [16], [17], we propose to use polling multiple access in this paper as in [11] in order to prevent collisions stemming from attempts from multiple sources to transmit at the same time. Moreover, the polling message can be sent in the form of a wake-up radio signal upon reception of which the scheduled user wakes up and carries out the sampling and transmission operations before going back to sleep [18], [19], [20], [21]. This prevents the unscheduled sources from wasting unnecessary energy, making it suitable for low power users in massive IoT networks. Fig. 1 illustrates Scenario A with N sources. Users may use different modulation and coding schemes depending on their distance from the BS, and therefore, they have heterogeneous packet service time (transmission time) requirements which we characterize with their first two moments in the current paper. Moreover, packet drops are assumed to occur according to heterogeneous packet error probabilities for each of the user-to-server transmissions, and these statistics are assumed to be a-priori known by the server.

The second application scenario (Scenario B) also worth of investigation is the wireless downlink scenario where a base station sends time-sensitive information to a number of users through unreliable channels in a wireless broadcast network [22] and we are interested in minimizing the weighted AoI maintained at the users. Wake-up radio can also accompany the system model of [22] for energy efficiency purposes. In contrast to the first scenario, it is always the server which is sending time-sensitive information on the downlink and the feedback channel from the users back to the server may not exist. Therefore, maintaining a replica of the AoI processes

at the server side may not even be possible, which is another reason for using age-agnostic scheduling. Both scenarios A and B can be addressed with the proposed methods of the current paper.

Two main approaches stand out in the literature for scheduling in multi-source status update systems. In age-aware scheduling, there are schedulers that make use of the framework of Markov decision processes (MDP) and dynamic programming (DP) with proven optimality [12], [23], [24]. However, computing the optimal policy using DP becomes prohibitive for large values of N since the state-space of the MDP grows exponentially with N . Hence, curse of dimensionality inherent to MDP models [25] hinders one from developing schedulers for massive connectivity scenarios. The alternative age-aware approach is the sub-optimal index-based scheduling such as the ones proposed in [22], [26], and [27], for which the scheduling policy is relatively easier to develop off-line. However, their runtime implementation still requires the server to continuously keep track of the AoI of the sources, and $O(N)$ computational complexity to choose the source with the highest index. Although far lower than the complexity of DP counterparts, the runtime computational burden of such algorithms is one of the limitations of age-aware schemes for massive IoT networks. Moreover, age-aware schemes fail to extend to scenarios in which the server cannot always perfectly maintain a replica of the actual AoI. This may happen since in the first scenario, monitors may be located away from the server introducing additional delays and losses in the forward and reverse channels from the server to monitors, and in the second scenario, ACK/NACK messages may not be sent from the users back to the server. In the alternative age-agnostic schemes, and in particular cyclic scheduling [28], [29], the server does not need to keep track of the ages of the individual sources while making scheduling decisions. Instead, the server develops a cyclic schedule, i.e., a fixed pattern of user transmissions which repeats itself, with an off-line algorithm using a-priori known statistical information. Once the schedule is developed, its runtime complexity is $O(1)$ which makes it an exceptionally suitable scheme for massive IoT networks.

The main scope of this paper is the off-line development of low complexity cyclic schedules for weighted AoI minimization that can scale to very large number of users encountered

in massive IoT settings. In this work, we extend our previous work on cyclic scheduling [29] in three directions: First, we provide an analytical method to compute the weighted AoI in the same setting as in [29] but also taking into consideration the case of heterogeneous packet errors, whereas no packet errors were assumed in [29]. Second, [29] shows in the absence of packet errors and for the specific case of two sources that, the optimal cyclic scheduler is of the form $(1, \Theta)$, which represents a cyclic schedule where one scheduling instance of one of the sources is followed by Θ scheduling instances of the other source. In this paper, we show that for two heterogeneous sources and in the presence of packet errors, a near-optimal cyclic scheduler is in the form of a mixture of $(1, \Theta)$ and $(1, \Theta + 1)$ cyclic schedulers, and we provide a novel algorithm for its construction. Although the $N = 2$ case is definitely not a large-scale scenario, having the optimum solution in hand for this special case enables us to compare various heuristic approaches against the optimum solution. Third, [29] proposes the Insertion Search (IS) heuristic algorithm for $N > 2$ but its computational complexity exceeding $O(N^5)$ makes it unfit for massive connectivity scenarios. In here, we propose a heuristic algorithm which is entirely different than IS and which can scale to massive IoT scenarios. A preliminary study on the construction of scalable cyclic schedulers can be found in [30].

Our main contributions are summarized as follows:

- We provide two novel analytical frameworks, MC-based and MGF-based, for the weighted AoI computation of a given cyclic schedule, both of which are substantially different than the method described in [29], since the behavior of the constructed cyclic schedule will no longer be cyclic due to packet drops. MC-based and MGF-based approaches both have their own merits which will become apparent in Section V.
- For the special case of $N = 2$, we prove that given any $\epsilon > 0$, our algorithm constructs a cyclic scheduler whose weighted AoI is within ϵ of the actual optimum. Therefore, we establish a lower bound for the weighted AoI for the $N = 2$ case, which can be used as a baseline for comparative evaluation of age-agnostic algorithms in this special case.
- We propose algorithms for the construction of a cyclic schedule appropriate for massive IoT scenarios with the goal of weighted AoI minimization, which are based on convex optimization and well-established packet spreading algorithms. We show through numerical results and simulations that our algorithms outperform existing age-agnostic multi-source schedulers in systems involving a very large number of users, in terms of weighted AoI minimization. Moreover, their relatively low off-line complexity as well as their $O(1)$ runtime complexity makes cyclic scheduling a good fit for massive connectivity scenarios.

The remainder of the paper is organized as follows. Section II overviews the related work. Section III presents the system model. Section IV develops the analytical method to obtain the weighted AoI when cyclic scheduling is employed. Section V derives the optimum cyclic scheduler in closed form for the case of two sources. Section VI presents the proposed scalable scheduling algorithm for general

number of sources. Section VII describes the existing baseline age-agnostic schemes along with the modifications we have introduced in this work. Section VIII presents extensive numerical results. Section IX concludes the paper.

II. RELATED WORK

AoI metric was first proposed in a single-source, single-server queuing scenario [5]. A survey of existing research on AoI and its applications can be found in [6], [7], and [8]. Several AoI stochastic analysis methods have been developed in the literature. For random arrivals (RA) where status updates take place according to a random process, e.g., such as the Poisson process, the stochastic hybrid systems (SHS) approach of [4] is generally used in the literature for obtaining the mean AoI, whereas SHS can also be used to obtain the moment generating function (MGF) and also the higher moments of AoI in certain scenarios [31]. The reference [32] derives the MGF of the AoI and peak AoI in a multi-source M/G/1/1 queuing model for several preemption policies. On the other hand, for multi-source GAW systems (and also for a special multi-source RA system with one single packet buffer), an alternative absorbing Markov chain (AMC) based method is proposed to find the distribution of AoI in [33], which requires a-priori information about the service time distributions. The current paper is different than [33], since we only focus on the average AoI in this paper (as opposed to its distribution), which only depends on the first two moments of the per-source service times (and not on their distributions), and also on the packet error probabilities.

Age-aware scheduling for AoI minimization has extensively been studied in the literature. Maximum age first (MAF) policy [34], [35], [36], where the source with the highest instantaneous age is scheduled, and maximum-age-difference drop (MAD) policy [37], where the source which would result in the maximum drop in age is scheduled, have been proposed in the literature for homogeneous multi-source settings. For heterogeneous scenarios, Max-Weight and Whittle-index policies [22], [26], [27] have been proposed that are shown to perform very well despite the lack of proof of optimality in general scenarios. Timeliness of communication in federated learning (FL), which was addressed in [38], is another domain where age-aware scheduling policies have been utilized. Further, [39] proposes an age-aware scheduling policy for FL which also takes into account the instantaneous channel qualities.

Age-aware scheduling policies require the scheduler to continuously track the age of the sources, and therefore, can introduce a significant communication overhead, especially in channels susceptible to packet drops. Moreover, in open-loop communication systems where feedback on packet drops is absent, age-aware scheduling may not even be feasible. Therefore, age-agnostic cyclic scheduling has recently emerged as a low cost open-loop solution compared to age-aware scheduling [28], [29], [40], [41]. A framework called *Eywa* was recently introduced in [28], where the goal is to construct *almost uniform cyclic schedulers (AUS)*, which is a special class of cyclic schedulers designed to distribute the scheduling instances of a given source as uniformly as possible within the cycle. *Eywa* works in a discrete time setting and assumes that all sources have the same deterministic service times with heterogeneous packet errors. However, in real systems, service

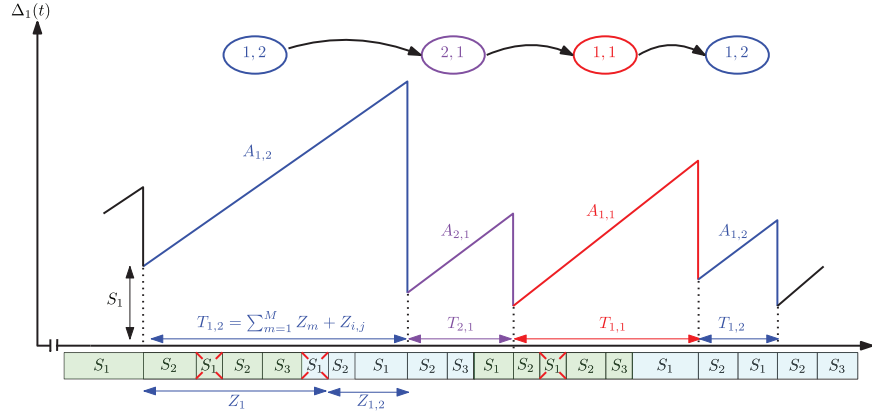


Fig. 2. Evolution of the AoI process of source-1 for $P = [1, 2, 1, 2, 3]$ showing a realization of the area $A_{1,2}$ when $M = 1$.

time requirements may be different across the sources, for which case reference [29] obtains the optimal cyclic scheduler that minimizes the weighted AoI for two heterogeneous sources, and also develops a heuristic scheduler for the case of several sources, but in the absence of packet errors. The focus of our current paper is on age-agnostic scheduling in a massive connectivity setting, and also in the presence of packet errors.

III. SYSTEM MODEL

Consider the status update system shown in Fig. 1, consisting of N sources indexed by $n = 1, 2, \dots, N$. The sources are served according to a cyclic schedule, where each source, once scheduled, samples its associated random process and generates a data packet which is transmitted through the shared channel. We consider a random delay channel which is prone to packet drops, where the n th source is subject to a packet drop probability $0 \leq p_n < 1$ and the packet success probability $u_n = 1 - p_n$. The random delay (service time) experienced by the l th ($l \geq 1$) packet of the n th source is denoted by $S_{n,l}$ which is an i.i.d. (independent and identically distributed) sequence with mean $s_n = \mathbb{E}[S_n]$, second moment $q_n = \mathbb{E}[S_n^2]$, variance $v_n = q_n - s_n^2$, and squared coefficient of variation (scov) $c_n = \frac{v_n}{s_n^2}$. The fixed wake-up signal transmission time for querying purposes is assumed to be much shorter compared to the information packet service times in both Scenarios A and B. Once a scheduled source has finished its transmission, a new sample from the next source in the schedule will be immediately sampled and transmitted. The transmitter (a user in the case of Scenario A and the server in the case of Scenario B) is assumed to lack information on whether a packet was successfully received or dropped, due to the open-loop nature of the system.

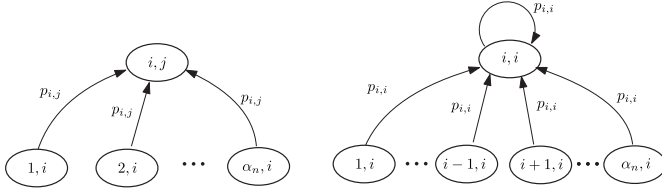
The cyclic schedule is characterized by a pattern $P = [P_0, P_1, \dots, P_{K-1}]$, which is a vector of size K such that $P_k \in \{1, 2, \dots, N\}$. Based on the pattern P , we generate cyclically left-shifted versions of this pattern denoted by $P^{(J)}$ where $0 \leq J \leq K-1$ denotes the number of shifts. As an example, $P^{(0)} = P$ and $P^{(1)} = [P_0^{(1)}, P_1^{(1)}, \dots, P_{K-1}^{(1)}] = [P_1, P_2, \dots, P_{K-1}, P_0]$. The cyclic scheduler selects source- $P_k^{(J)}$ for transmission at the scheduling instant $k + iK$, $0 \leq k < K$, $i = 1, 2, \dots$, where J is a discrete uniformly distributed random variable in the interval $[0, K-1]$. Therefore, the system starts operation from a uniformly random position in the given pattern P at the

initial scheduling instance $k = 0$. As an example, let $N = 3$, $P = [1, 2, 1, 2, 3]$ and $J = 1$. In this case, $P^{(1)} = [2, 1, 2, 3, 1]$ and the cyclic scheduler will transmit the following sequence of source packets starting from the initial scheduling instance: 2, 1, 2, 3, 1, 2, 1, 2, 3, 1, 2, 1, 2, 3, ... A pattern is said to be feasible if each of the N sources appears at least once in the pattern since otherwise AoI of one or more sources would be unbounded. Thus, the pattern size should be larger than or equal to the number of users, i.e., $K \geq N$. Stemming from the choice of the random initial scheduled source, the resulting AoI process for source- n observed at the server, denoted by $\Delta_n(t)$, $1 \leq n \leq N$, is an asymptotically stationary and mean-ergodic process even in the scenario when all the service times are deterministic and all error probabilities are zero. Actually, $\Delta_n(t)$ increases linearly with unit slope in time and upon successful reception (without packet drop) of a data sample from source- n , the age process drops to the value of the delay experienced by the particular data packet that was successfully received. Let Δ_n denote the steady-state random variable for the associated age process $\Delta_n(t)$ and let $\mathbb{E}[\Delta_n]$ denote its mean which is also the time average of the age process $\Delta_n(t)$ due to its ergodicity. Fig. 2 depicts the evolution of the age process of source-1 for the above pattern P , where failed source-1 transmissions are crossed out in red. The per-source service times considered in Fig. 2 are heterogeneous and random. As shown in Fig. 2, $\Delta_1(t)$ continues to increase linearly during the transmission of samples from other sources as well as during failed source-1 transmissions, and drops upon successful reception of a source-1 sample. We define the weighted AoI as the mean of the random variable $\Delta = \sum_{n=1}^N w_n \Delta_n$, i.e.,

$$\mathbb{E}[\Delta] = \sum_{n=1}^N w_n \mathbb{E}[\Delta_n], \quad (1)$$

where the normalized source weights w_n , $\sum_{n=1}^N w_n = 1$, are used to prioritize the sources.

The main goal of this work is to design a cyclic scheduler, i.e., a transmission pattern P , with an attempt to minimize the weighted AoI by taking packet drop probabilities into consideration. As the first step towards realizing this goal, we present two approaches for analytically obtaining the weighted AoI of a given transmission pattern.

Fig. 3. Partial state transition diagram for $\alpha_n > 2$.

IV. AOI ANALYSIS

Even though we start off with a cyclic schedule, due to packet drops, the actual schedule of successful transmissions will no longer behave according to the constructed schedule. Thus, weighted AoI computation is a difficult task even for two sources. In this section, we present two methods, namely MC- and MGF-based methods, which can be used to compute the average AoI of a cyclic scheduling scheme. Each method has its merits, where the former one will be used to construct a near-optimal scheduler for two heterogeneous sources and the latter one will be employed in the design of large-scale cyclic schedulers.

To derive the AoI expression in here, we first consider a single source (say source- n) and characterize its AoI process using a Markov chain (MC) formulation. Let $\alpha_n, n = 1, \dots, N$ denote the number of times that source- n appears in the pattern P . Note that $\alpha_n > 0$, for all n , since otherwise it will not be a feasible schedule (weighted AoI will be unbounded). There are α_n locations within the pattern P where the AoI of source- n can drop. Therefore, based on where two consecutive AoI drops occur relative to the pattern P , we can define α_n^2 states for the MC. Let (i, j) for $i, j \in \{1, 2, \dots, \alpha_n\}$ denote an AoI cycle starting from the i th scheduling instance of source- n and ending in the j th scheduling instance of source- n . Note that (i, j) represents the i th and the j th scheduling instances of source- n relative to the pattern P . For example, if $P = [1, 2, 1, 2, 1, 3]$, then our source transmissions will follow the following schedule: 1, 2, 1, 2, 1, 3, 1, 2, 1, 2, 1, 3, ... Now, consider the AoI process of source-1. If an AoI drop occurs after the first scheduling instance of source-1 in the first pattern period and the next AoI drop occurs at the end of the third scheduling instance of source-1 in the second pattern period, then this AoI cycle would belong to state $(1, 3)$. Each state transition occurs after successfully transmitting a sample without packet drops from source- n . Fig. 3 shows a partial state transition diagram for the MC.

As shown in Fig. 3, state (i, j) is only directly accessible by states of the form (k, i) and they all share the same transition probability $p_{i,j}$. Only states of the form (i, i) have self-transitions. The transition probabilities are given as,

$$p_{i,j} = \frac{(1 - p_n)p_n^{j-i-1}}{1 - p_n^{\alpha_n}}, \quad \text{if } j > i, \quad (2)$$

$$p_{i,i} = \frac{(1 - p_n)p_n^{j+\alpha_n-i-1}}{1 - p_n^{\alpha_n}}, \quad \text{if } i \geq j. \quad (3)$$

To understand the above transition probabilities and state-space of the MC, let us consider the following pattern $P = [1, 2, 1, 2, 1, 3]$ (note that this pattern is slightly different from the one in Fig. 2). Suppose we are interested in finding the AoI of source-1. For this pattern, $\alpha_1 = 3$ and therefore we would

have 9 states for the MC. Based on which scheduling instances the successive AoI drops occur for source-1, the MC will have the states $(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2)$ and $(3, 3)$. For example, state $(3, 2)$ refers to an AoI cycle with the first drop occurring at the third scheduling instance of source-1 and the next drop incurred at the second scheduling instance of source-1. To better understand these states, we refer to Fig. 2. Now, suppose we want to find the probability of transitioning from state $(1, 1)$ to $(1, 3)$. If this transition happens in the first pattern period (cycle), it means that our next source-1 transmission was unsuccessful and the second source-1 transmission was successful which would occur with probability $p_1(1 - p_1)$. If the transition takes place in the second period, then this would mean that our next four source-1 transmissions failed and we were successful at the fifth transmission which would happen with probability $p_1^4(1 - p_1)$. Similarly, this would happen in the third period with probability $p_1^7(1 - p_1)$ and in the k th period with probability $p_1^{3(k-1)}p_1(1 - p_1)$. Therefore, the resulting probability that we transition from state $(1, 1)$ to $(1, 3)$ is $\sum_{k=1}^{\infty} p_1(1 - p_1)p_1^{3(k-1)}$ which gives the transition probability as $\frac{(1-p_1)p_1}{1-p_1^3}$ which exactly aligns with (2). Derivation of (3) follows a similar analysis.

Since these are finite-state MCs with no absorbing states for $p_n > 0$, they are positive recurrent and irreducible. Since they contain self-loops, they are aperiodic and hence are ergodic. Let $\pi = \{\pi_{(i,j)}\}$ for $i, j \in \{1, 2, \dots, \alpha_n\}$ denote the stationary distribution of the MC. Then, the stationary distribution takes the following form,

$$\pi_{i,j} = \frac{p_{i,j}}{\alpha_n}. \quad (4)$$

To compute the average AoI, we partition the AoI graph based on the states of the MC.

Theorem 1: Let $A_{i,j}$ denote the area under the age curve when the two consecutive age drops are placed at the i th and j th scheduling instances of the cycle. Let $T_{i,j}$ denote the time spent during the aforementioned age drops. Then, the average AoI of source- n is given by,

$$\mathbb{E}[\Delta_n] = \frac{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \pi_{i,j} \mathbb{E}[A_{i,j}]}{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \pi_{i,j} \mathbb{E}[T_{i,j}]}. \quad (5)$$

The proof of Theorem 1 is given in Appendix IX. As depicted in Theorem 1, to compute the average AoI, we need to find $\mathbb{E}[A_{i,j}]$ and $\mathbb{E}[T_{i,j}]$. $A_{i,j}$ consists of the segments of the age curve with consecutive age drops happening at the i th and j th scheduling instances of source- n . Suppose an AoI drop happens at the i th scheduling instance, then the next AoI drop may happen at the j th scheduling instance after going through multiple rounds of the entire pattern P . This is illustrated in Fig. 2 considering the average AoI of source-1. Let M denote the number of rounds of the entire pattern P that has elapsed before the next successful transmission of a source- n sample occurring at the j th scheduling instance. For example, if the next successful transmission occurs in the second round, then $M = 1$. Then, $M = \hat{M} - 1$, where $\hat{M}$ is a geometric random variable with parameter $p_n^{\alpha_n}$, i.e., $\hat{M} \sim \text{Geom}(p_n^{\alpha_n})$. Let $Z_{i,j}$ denote the time duration elapsed starting from the beginning of the $(M + 1)$ th round to the AoI drop occurring at the

j th scheduling instance of source- n . Let Z_m denote the time duration of the m th round. Then, $\mathbb{E}[A_{i,j}]$ can be written as,

$$\mathbb{E}[A_{i,j}] = \frac{\mathbb{E}\left[\left(2S_n + \sum_{m=1}^M Z_m + Z_{i,j}\right)\left(\sum_{m=1}^M Z_m + Z_{i,j}\right)\right]}{2} \quad (6)$$

$$= \frac{\tilde{m}\hat{s}^2 + \bar{m}(2s_n\hat{s} + \hat{v}) + \tilde{z}_{i,j} + 2\bar{z}_{i,j}(\bar{m}\hat{s} + s_n)}{2}, \quad (7)$$

where $\hat{s} = \mathbb{E}[Z_m] = \sum_{k=1}^N \alpha_k s_k$, $\hat{v} = \sum_{k=1}^N \alpha_k v_k$, $\mathbb{E}[M] = \bar{m}$, $\mathbb{E}[M^2] = \tilde{m}$, $\mathbb{E}[Z_{i,j}] = \tilde{z}_{i,j}$ and $\mathbb{E}[Z_{i,j}^2] = \tilde{z}_{i,j}^2$. Similarly, $\mathbb{E}[T_{i,j}]$ can be written as,

$$\mathbb{E}[T_{i,j}] = \mathbb{E}\left[\sum_{m=1}^M Z_m + Z_{i,j}\right] = \bar{m}\hat{s} + \bar{z}_{i,j}. \quad (8)$$

Then, by substituting (7) and (8) into (5), we find the average AoI of source- n as follows,

$$\mathbb{E}[\Delta_n] = \frac{\tilde{m}\hat{s}^2 + \bar{m}(2s_n\hat{s} + \hat{v}) + \sum_{i,j} \pi_{i,j} \tilde{z}_{i,j}}{2(\bar{m}\hat{s} + \sum_{i,j} \pi_{i,j} \bar{z}_{i,j})} + \frac{(\bar{m}\hat{s} + s_n) \sum_{i,j} \pi_{i,j} \tilde{z}_{i,j}}{\bar{m}\hat{s} + \sum_{i,j} \pi_{i,j} \bar{z}_{i,j}}. \quad (9)$$

An alternative method to obtain the weighted AoI of a cyclic schedule is through the use of total law of expectation in conjunction with moment generating functions (MGF). The motivation for development of this alternative method is due to the fact that the age expression obtained by this method can be easily fashioned into an optimization problem that can be used to construct well-performing heuristic solutions for the development of cyclic schedules for a large number of sources. For this purpose, let $S_{n,l}$ denote the service time of the l th successful transmission of source- n and let $\tilde{S}_{n,l}$ denote the time duration between the end of the l th successful transmission and the beginning of the $(l+1)$ th successful transmission of source- n . Note that $\tilde{S}_{n,l}$ is the sum of the service times of all packets generated from sources other than source- n between two consecutive successful transmissions of source- n , denoted by $\tilde{S}_{n,l}^o$, and also of the service times of failed packets belonging to source- n itself, denoted by $\tilde{S}_{n,l}^u$. Clearly, $\tilde{S}_{n,l} = \tilde{S}_{n,l}^o + \tilde{S}_{n,l}^u$. Let the random variable $\tilde{S}_n$ (steady-state random variable associated with the random process $\tilde{S}_{n,k}$ as $k \rightarrow \infty$) have mean $\tilde{s}_n = \mathbb{E}[\tilde{S}_n]$, second moment $\tilde{q}_n = \mathbb{E}[\tilde{S}_n^2]$, variance $\tilde{v}_n = \tilde{q}_n - \tilde{s}_n^2$ and $\text{scov } \tilde{c}_n = \frac{\tilde{v}_n}{\tilde{s}_n}$. Based on [29], the mean AoI for source- n can be written in terms of the first two moments of the random variables S_n and $\tilde{S}_n$ as,

$$\mathbb{E}[\Delta_n] = \frac{2s_n^2 + 4s_n\tilde{s}_n + q_n + \tilde{q}_n}{2(s_n + \tilde{s}_n)} \quad (10)$$

$$= \frac{s_n^2(c_n + 3) + \tilde{s}_n^2(\tilde{c}_n + 1) + 4s_n\tilde{s}_n}{2(s_n + \tilde{s}_n)}. \quad (11)$$

Therefore, one needs to compute the parameters $\tilde{s}_n$ and $\tilde{q}_n$ (or $\tilde{c}_n$) in order to find the average AoI of source- n from (10) for which we present an MGF-based method in Appendix IX.

V. TWO SOURCE CYCLIC SCHEDULERS ($N = 2$)

In this section, we analyze the average weighted age expression obtained by the MC formulation for two heterogeneous

sources, and present algorithms to construct two-source cyclic schedulers which we subsequently prove to be near optimal.

Remark 1: Throughout this paper, when we refer to a schedule as optimal, it simply means that it is a cyclic scheduler that achieves the lowest weighted AoI among all possible cyclic schedulers.

A. Near Optimal Two Source Scheduler (NOTS)

Let $\mathbf{r} = \{r_1, r_2, \dots, r_{\alpha_1}\}$ represent the placement vector of the schedule with respect to source-1, where r_i is the number of source-2 scheduling instances between the i th and the $(i+1)$ th scheduling instances of source-1. Note that $\alpha_2 = \sum_i r_i$. Any scheduling pattern P corresponding to a two source cyclic scheduler, can be uniquely represented by this placement vector $\mathbf{r}$. For example, the pattern $P = [1, 2, 2, 2, 1, 2, 1, 2]$ would correspond to placement vector $\mathbf{r} = \{3, 1, 1\}$. Let $a = \frac{\alpha_2}{\alpha_1}$, $s = as_2 + s_1$ and $v = av_2 + v_1$. Then, Theorem 2 gives the expression for the time averaged AoI of source-1.

Theorem 2: Let $\gamma = \{r_1, r_2, \dots, r_{\alpha_1}, r_1, r_2, \dots, r_{\alpha_1}\}$. Then, the $\mathbb{E}[\Delta_1]$ will be given by,

$$\mathbb{E}[\Delta_1] = \frac{(1+p_1)}{2(1-p_1)}s + \frac{v}{2s} + s_1 + s_2^2 \frac{(1-p_1)^2}{2s\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} (\tilde{r}(i) - \alpha_1 a^2 i^2) p_1^{i-1}, \quad (12)$$

where $\tilde{r}(i)$ is defined as,

$$\tilde{r}(i) = \sum_{j=1}^{\alpha_1} \left(\sum_{k=j}^{j+i-1} \gamma_k \right)^2. \quad (13)$$

Corollary 1: Let $\mathbf{l} = \{l_1, l_2, \dots, l_{\alpha_2}\}$ represent the placement vector of the schedule with respect to source-2 and let $\zeta = \{l_1, l_2, \dots, l_{\alpha_2}, l_1, l_2, \dots, l_{\alpha_2}\}$. Then, the average AoI of source-2, $\mathbb{E}[\Delta_2]$, will be given by,

$$\mathbb{E}[\Delta_2] = \frac{(1+p_2)}{2a(1-p_2)}s + \frac{v}{2s} + s_2 + s_1^2 \frac{a(1-p_2)^2}{2s\alpha_2(1-p_2^{\alpha_2})} \sum_{i=1}^{\alpha_2} \left(\tilde{l}(i) - \frac{\alpha_2 i^2}{a^2} \right) p_2^{i-1}, \quad (14)$$

where $\tilde{l}(i)$ is defined as,

$$\tilde{l}(i) = \sum_{j=1}^{\alpha_2} \left(\sum_{k=j}^{j+i-1} \zeta_k \right)^2. \quad (15)$$

Remark 2: When $p_1 = p_2 = 0$, equations (12) and (14) reduce to the average AoI expression obtained in [29] which is a special case of the model studied here.

The proofs of Theorem 2 and Corollary 1 are presented in Appendix IX. Next, to construct the optimal cyclic scheduler, we need to find the optimal placement vector $\mathbf{r}^*$ that minimizes weighted AoI, $w_1\mathbb{E}[\Delta_1] + w_2\mathbb{E}[\Delta_2]$. To proceed with our analysis, for fixed α_1 and α_2 , we first find the optimal placement vector that minimizes (12).

Theorem 3: For fixed α_1 and α_2 , let r_i^* represent the elements of the optimal placement vector that minimizes (12). Then, $r_i^* \in \{[a], [a]\}$ and satisfies the following equations,

$$\sum_{i=1}^{\alpha_1} \mathbf{1}_{\{r_i^*=[a]\}} = \alpha_1(1 + a - [a]), \quad (16)$$

$$\sum_{i=1}^{\alpha_1} \mathbf{1}_{\{r_i^* = \lfloor a \rfloor\}} = \alpha_1(1 + \lfloor a \rfloor - a). \quad (17)$$

Moreover, the optimal placement vector $\mathbf{r}^*$ can be constructed by distributing the $\lceil a \rceil$ and $\lfloor a \rfloor$ terms as uniformly as possible within the placement vector.

Algorithm 1 Uniform Arrangement of Placement Vectors

Input: α_1, α_2

```

 $a = \frac{\alpha_2}{\alpha_1}, b_1 = \lfloor a \rfloor, b_2 = \lceil a \rceil$ 
 $c_1 = \alpha_1(\lceil a \rceil - a), c_2 = \alpha_1(1 + a - \lceil a \rceil)$ 
while  $\min(c_1, c_2) > 1$  do
  if  $c_1 > c_2$  then
     $c_1 \leftarrow c_2, c_2 \leftarrow c_1$ 
     $b_1 \leftarrow b_2, b_2 \leftarrow b_1$ 
  end if
   $c_0 = \frac{c_2}{c_1}$ 
   $b_1 \leftarrow \{b_1, b_2 \times \lfloor c_0 \rfloor\}, b_2 \leftarrow \{b_1, b_2 \times \lceil c_0 \rceil\}$ 
   $c_1 \leftarrow c_1(\lceil c_0 \rceil - c_0), c_2 \leftarrow c_1(1 + c_0 - \lceil c_0 \rceil)$ 
end while

```

The proof of Theorem 3 is given in Appendix IX. In Theorem 3, (16) and (17) simply mean that the number of elements in optimal placement vector $\mathbf{r}^*$ with value $\lceil a \rceil$ is $\alpha_1(1 + a - \lceil a \rceil)$ and the number of elements with value $\lfloor a \rfloor$ is $\alpha_1(1 + \lfloor a \rfloor - a)$. As seen by Theorem 3, to construct the optimal placement vector, we need to spread the $\lceil a \rceil$ and $\lfloor a \rfloor$ terms as uniformly as possible within the placement vector. This is one of the key differences from the work in [29] where the ordering of the placement vector is inconsequential. The optimal placement vector can be constructed by hierarchically spreading different sub-blocks of the placement vector as given by Algorithm 1.

In Algorithm 1, $b_2 \leftarrow \{b_1, b_2 \times \lceil c_0 \rceil\}$ implies that the new b_1 is an array consisting of one instance of b_1 and $\lceil c_0 \rceil$ instances of b_2 . As an application of Algorithm 1, let us consider a scenario with $\alpha_2 = 41$ and $\alpha_1 = 11$. One possible placement vector for this case is $\mathbf{r} = \{3, 3, 3, 4, 4, 4, 4, 4, 4, 4\}$ which is a vector of 3 threes and 8 fours. Now, we need to uniformly arrange this placement vector. For this purpose, we try to uniformly distribute the 8 fours among the 3 threes. This will give us two blocks of $\{3, 4, 4, 4\}$ and one block of $\{3, 4, 4\}$. Next, we need to spread these sub-blocks as uniformly as possible. Since the minimum of the number of instances of these two blocks is one, the algorithm stops at this stage and returns $\mathbf{r} = \{3, 4, 4, 4, 3, 4, 4, 4, 3, 4, 4\}$ as the optimal placement vector. The optimal placement vector generated by Algorithm 1 is always a mixture of placement vectors of the form $\mathbf{r} = \{\Theta\}$ and $\mathbf{r} = \{\Theta + 1\}$, where $\Theta = 3$ in this example.

Note that the above treatment tries to minimize $\mathbb{E}[\Delta_1]$ by scheduling source-1 transmissions as uniformly as possible for a fixed α_1 and α_2 . When we uniformly distribute source-1 transmissions, we also notice that at the same time, it allows us to uniformly distribute the S_2 as well. Hence, the above structure of the placement vector jointly minimizes both $\mathbb{E}[\Delta_1]$ and $\mathbb{E}[\Delta_2]$. This is one of the similarities with the work in [29]. Therefore, what remains is to find the best possible α_1 and α_2 . This can be reduced to finding the best possible rational number for a that minimizes the weighted AoI.

To find the optimal a , we first find the average AoI of the round robin (RR) policy, denoted by $\mathbb{E}[\Delta_{RR}]$, which can be obtained by setting $\alpha_1 = \alpha_2 = a = 1$ in (12) and (14),

$$\begin{aligned} \mathbb{E}[\Delta_{RR}] &= \frac{(s_1 + s_2)}{2} \left(w_1 \frac{(1 + p_1)}{(1 - p_1)} + w_2 \frac{(1 + p_2)}{(1 - p_2)} \right) \\ &\quad + \frac{v_1 + v_2}{2(s_1 + s_2)} + s_1 w_1 + s_2 w_2. \end{aligned} \quad (18)$$

Since the first term in $\mathbb{E}[\Delta_1]$ is linear in a , there will be an a for which the weighted average AoI of source-1 will be greater than that of the RR policy. We do not have to consider the values of a beyond this point since then the AoI performance of the RR policy will be superior. Therefore, we only have to consider the values of a for which the term $w_1 \mathbb{E}[\Delta_1]$ is less than the weighted AoI of the RR policy. Let a_{max} be the smallest a such that $w_1 \mathbb{E}[\Delta_1] > \mathbb{E}[\Delta_{RR}]$. Similarly, we note that the first term in $\mathbb{E}[\Delta_2]$ is linear in $\frac{1}{a}$. Therefore, we can find a a_{min} value, which is the largest a such that $w_2 \mathbb{E}[\Delta_2] > \mathbb{E}[\Delta_{RR}]$. Hence, the search space of a is bounded by a_{min} and a_{max} .

Algorithm 2 Algorithm to Find the Near-Optimal Cyclic Pattern

Input: $\alpha \geq \alpha_\epsilon$ given by Theorem 4

$\alpha_1 = \alpha, \alpha_2 = \alpha, AoI_{min} = \mathbb{E}[\Delta_{RR}], \mathbf{r}^* = \{1\}$

$a_{max} = \inf\{a : w_1 \mathbb{E}[\Delta_1] > \mathbb{E}[\Delta_{RR}]\}$

```

while  $\frac{\alpha_2}{\alpha_1} < a_{max}$  do
   $(\tilde{\alpha}_1, \tilde{\alpha}_2) = \text{Coprime}(\alpha_1, \alpha_2)$ 
   $\mathbf{r} = \text{ALG 1}(\tilde{\alpha}_1, \tilde{\alpha}_2)$ 
   $AoI = w_1 \mathbb{E}[\Delta_1'] + w_2 \mathbb{E}[\Delta_2']$ 
  if  $AoI_{min} > AoI$  then
     $\mathbf{r}^* = \mathbf{r}, AoI_{min} = AoI$ 
  end if
   $\alpha_2 = \alpha_2 + 1$ 
end while

```

$\alpha_1 = \alpha, \alpha_2 = \alpha$

$a_{min} = \sup\{a : w_2 \mathbb{E}[\Delta_2] > \mathbb{E}[\Delta_{RR}]\}$

```

while  $\frac{\alpha_2}{\alpha_1} > a_{min}$  do
   $(\tilde{\alpha}_1, \tilde{\alpha}_2) = \text{Coprime}(\alpha_1, \alpha_2)$ 
   $\mathbf{r} = \text{ALG 1}(\tilde{\alpha}_1, \tilde{\alpha}_2)$ 
   $AoI = w_1 \mathbb{E}[\Delta_1'] + w_2 \mathbb{E}[\Delta_2']$ 
  if  $AoI_{min} > AoI$  then
     $\mathbf{r}^* = \mathbf{r}, AoI_{min} = AoI$ 
  end if
   $\alpha_1 = \alpha_1 + 1$ 
end while

```

Output: $\mathbf{r}^*$

From this point onwards, we represent a two source cyclic schedule only using the tuple (α_1, α_2) where the optimal placement vector is found using Algorithm 1. For a fixed a , consider the patterns (α_1, α_2) and $(k\alpha_1, k\alpha_2)$ where $k \in \mathbb{N}$. If $\mathbf{r}^*$ is the optimal placement vector for (α_1, α_2) , by simply repeating $\mathbf{r}^*$, k times, we can obtain the optimal placement vector for the pattern $(k\alpha_1, k\alpha_2)$ which yields the same average AoI. Therefore, in the bounded region for a , we only need to compare rationals in their simplest form. Algorithm 2 can be used to find a near-optimal a and $\mathbf{r}^*$ which minimize the weighted AoI. In Algorithm 2, $\mathbb{E}[\Delta_i']$ represents the AoI of the i th source with respect to the placement vector of source-1,

ALG 1 ($\tilde{\alpha}_1, \tilde{\alpha}_2$) is the output of Algorithm 1 for the selected $\tilde{\alpha}_1, \tilde{\alpha}_2$ and **Coprime**(α_1, α_2) returns the co-primes of α_1 and α_2 . Even though the average AoI depends on both a and α_1 , we show in Theorem 4 that fixing α_1 to a large value and finding a using Algorithm 2 is sufficient to be as close as desired to the optimal cyclic schedule. The proof of Theorem 4 is given in Appendix IX.

Theorem 4: Let the weighted AoI of the optimal cyclic scheduler be denoted by Δ^* and the weighted AoI of the cyclic schedule constructed by Algorithm 2 with parameter α be denoted by Δ^α . Then, $\forall \epsilon > 0, \exists \alpha_\epsilon \in \mathbb{Z}$ such that $|\Delta^\alpha - \Delta^*| < \epsilon, \forall \alpha \geq \alpha_\epsilon$.

Algorithm 3 Pseudo-Code to Refine the Schedule Using Sub-Patterns

Step 1: Set $\alpha \geq \alpha_\epsilon$ and find the near optimal a^* and α_1^* from Algorithm 2

Step 2: Set $\alpha_1 = \alpha_1^*$ and $\alpha_2 = a^* \alpha_1^*$ in Algorithm 1. Let SB be the set of all sub-blocks constructed across all stages in Algorithm 1.

Step 3: Find the sub-block in SB with the lowest weighted AoI.

B. Sub-Pattern Refinement

Algorithm 2 provides a means for finding a cyclic schedule as almost as good as the optimal cyclic schedule. However, it may end up generating long placement vectors to be as close as possible to the optimal. For example, suppose the optimal pattern is known to be $\{2, 3, 3\}$ where optimal $\alpha_1 = 3$. However, if we set $\alpha = 101$ in Algorithm 2, it would never output the pattern $\{2, 3, 3\}$ since α is not divisible by 3. Instead, it would output a pattern like $\{2, 3, 3, 3, 2, 3, 3, 2, 3, 3, \dots, 2, 3, 3\}$, where the sub-pattern $\{2, 3, 3, 3\}$ is followed by multiple repetitions of the sub-pattern $\{2, 3, 3\}$ so that the average AoI would be as close as possible to the optimal. Thus, we propose to further check if the sub-patterns existing within the output of Algorithm 2, could be a better solution. Algorithm 3 provides a method to further refine the cyclic schedule using these sub-patterns to produce a better and a simpler schedule.

VI. MULTI-SOURCE CYCLIC SCHEDULERS ($N > 2$)

We have outlined the structure followed by a near optimal scheduler for two heterogeneous sources in Section V. Even though the algorithms used to construct the two source scheduler can be proven to be near optimal, they cannot be easily extended to handle a larger number of sources ($N > 2$). In this section, we propose an algorithm to construct cyclic schedulers for a large number of sources. The algorithms used here are derived using the structure of the age expression in (11) presented in our MGF formulation.

A. Scalable Weighted AoI Minimization Scheduler (SAMS)

Let τ_n , $0 < \tau_n < 1$, $n \in \{1, \dots, N\}$, denote the link utilization of source- n , i.e., long-term fraction of time that the wireless link is occupied with the transmission of packets (successful or unsuccessful) from source- n . Since we focus on work-conserving servers, we have $\sum_{n=1}^N \tau_n = 1$. Moreover,

let the random variable $\tilde{S}_n^u$ (steady-state random variable associated with the random process $\tilde{S}_{n,k}^u$ as $k \rightarrow \infty$) have mean $\tilde{S}_n^u = \mathbb{E}[\tilde{S}_n^u]$. Then, τ_n takes the following form,

$$\tau_n = \frac{s_n + s_n^u}{s_n + \tilde{s}_n}. \quad (19)$$

For every successful source- n transmission, we must have on average $\frac{p_n}{u_n}$ unsuccessful transmissions from the same source. Therefore, we have $s_n^u = \frac{p_n}{u_n} s_n$, from which we obtain $\tilde{s}_n$ in terms of τ_n as,

$$\tilde{s}_n = \frac{s_n}{u_n \tau_n} - s_n. \quad (20)$$

By substituting (20) into (11), we obtain,

$$2 w_n \mathbb{E}[\Delta_n] = \underbrace{w_n s_n u_n (c_n + \tilde{c}_n)}_{a_n} \tau_n + \underbrace{\frac{w_n s_n (1 + \tilde{c}_n)}{u_n}}_{b_n} \frac{1}{\tau_n} + 2 w_n s_n (1 - \tilde{c}_n). \quad (21)$$

From (21), we can obtain the weighted AoI as,

$$\mathbb{E}[\Delta] = \sum_{n=1}^N \frac{a_n}{2} \tau_n + \frac{b_n}{2\tau_n} + \sum_{n=1}^N w_n s_n (1 - \tilde{c}_n). \quad (22)$$

If the $\tilde{c}_n$ values are given, then the minimization of $\mathbb{E}[\Delta]$ subject to $\sum_n \tau_n = 1$ is a convex optimization problem whose solution τ_n^* for $1 \leq n \leq N$ is given in Appendix IX. Let f_n^* denote the optimum long-term frequency of packet transmissions from source- n relative to the frequency of all transmissions. Actually, f_n^* is the long-term probability of a transmission belonging to source- n when the optimum utilization parameter τ_n^* is used for source- n . Since $f_n^* \propto \tau_n^*/s_n$ and $\sum_n f_n^* = 1$, the following closed-form expression holds for f_n^* ,

$$f_n^* = \frac{\tau_n^*}{s_n} \left(\sum_{m=1}^N \frac{\tau_m^*}{s_m} \right)^{-1}, \quad 1 \leq n \leq N. \quad (23)$$

The source utilization variables τ_n affect directly the parameters $\tilde{s}_n$ according to (20), but the $\tilde{c}_n$ values given in (11) are complex functions of the constructed schedule in addition to the choice of τ_n values. Therefore, it is difficult to pose a joint optimization problem involving both $\tilde{s}_n$ and $\tilde{c}_n$ for $1 \leq n \leq N$. We can also write (21) by collecting the terms involving $\tilde{c}_n$ as,

$$2 w_n \mathbb{E}[\Delta_n] = \frac{(u_n \tau_n - 1)^2}{u_n \tau_n} w_n s_n \tilde{c}_n + w_n s_n \left(c_n u_n \tau_n + \frac{1}{u_n \tau_n} + 2 \right). \quad (24)$$

From (24), we note that, for a fixed choice of τ_n values, the best choice for the minimization of weighted AoI would be to schedule the sources in such a way that $\tilde{c}_n$ parameters are as close to zero as possible, i.e., close to periodic transmissions of successful source- n packets. However, this may not always be possible even in the case of deterministic service times. Even if source- n transmissions are periodic, the time interval between two consecutive successful source- n transmissions may not necessarily be deterministic, i.e., $\tilde{c}_n$ will be non-zero, which stems from non-zero packet drop probability for source- n . Moreover, it may not be possible for source- n transmissions to be periodic due to constraints from other sources. When the service times are random, this situation becomes even more challenging.

Algorithm 4 Pseudo-Code for Obtaining K_n and K

Input: $f_n^*, 1 \leq n \leq N, \varepsilon \geq 0$;
Output: K_n, K ;
Step 1: Find $f_{\min}^* = \min_n f_n^*$;
Step 2: Set $K = \lceil \frac{f_{\min}^*}{f_{\min}^*} \rceil$;
Step 3: Find $R = \sum_{n=1}^N \lfloor K f_n \rfloor$;
Step 4: Sort all the sources in descending order according to the fractional part of $K f_n$ and set $K_n = \lceil K f_n \rceil$ for the top $K - R$ sources in this ordered list, otherwise set $K_n = \lfloor K f_n \rfloor$.

1) *Cyclic Schedule Construction:* Next, we present how to construct a well-performing cyclic scheduler based on the transmission frequencies computed above. As the next step for cyclic scheduling, we first need to write $f_n^* \approx K_n/K$ for integer K_n, K where K_n would represent the number of occurrences of source- n within the pattern P . For this, we propose Algorithm 4 in terms of the algorithm parameter $\varepsilon \geq 0$. The case of $\varepsilon = 0$ ensures that $K_n \geq 1, \forall n$, and has a small pattern size K . When ε increases, obviously the approximation K_n/K of the optimum frequency f_n^* improves but at the expense of increased K , and hence, increased storage requirements. The use of $\varepsilon > 0$ was motivated by the fact that NOTS produced optimal patterns that were a mixture of two sub-patterns. Therefore, by allowing the minimum number of source scheduling instances to be greater than 1, we make our policy more flexible and closer to NOTS when $N = 2$. The computational complexity of Algorithm 4 is $O(N \log N)$ due to the required sorting.

Next, we need a packet spreading algorithm which generates a transmission pattern in which source- n appears K_n times with pattern size $K = \sum_{n=1}^N K_n$, and all these appearances are as evenly spread as possible throughout the pattern. This problem with many variations had been studied in the context of *internet scheduling* where the goal was to share the link bandwidth fairly among multiple flows carrying fixed-size or variable-size packets; see [42] and the references therein for a collection of research papers on fair link sharing.

The packet spreading algorithm we propose to use in the present paper is based on the *deficit round robin* (DRR) algorithm proposed in [42] which has been successfully used in commercial routers due to its low computational complexity. DRR consists of rounds at each of which the deficit counters of each flow are incremented by the product of the so-called quantum and the weight of the flow. Subsequently, all the head-of-line packets waiting in the queue of each flow whose total packet size in bytes does not exceed the corresponding deficit counter, are served. The number of bytes that are served for a flow is then subtracted from the corresponding deficit counter, leaving a so-called deficit. In this way, multiple flows can be served in the same round. In our proposed spreading algorithm, our goal is to find a cyclic scheduler that spreads out source- n transmissions as much as possible while source- n appears K_n times in the pattern. Initially, all deficit counters, denoted by $B_n(t), 1 \leq n \leq N$, are set to zero. We modify the original DRR algorithm by allowing the value of the quantum change between rounds, which is slightly different than the original DRR scheduler [42]. Particularly, the quantum is chosen to ensure that one source is guaranteed to be inserted into the pattern at a given round while leaving zero deficit. The

pseudo-code for the proposed packet spreading algorithm is given in Algorithm 5 with computational complexity $O(NK)$. Special characteristics of the packet spreading are highlighted in Appendix X.

Algorithm 5 Algorithm for Constructing the Pattern P

Input: $N \geq 2, K_n, n = 1, \dots, N, K = \sum_{n=1}^N K_n$;
Output: Pattern P of size K ;
 $B_n(t) \leftarrow 0$; ($B_n(t)$: deficit counter for source- n)
for $k = 0$ **to** $K - 1$ **do**
 $m \leftarrow \arg \min_{1 \leq n \leq N} \frac{(1 - B_n(t))K}{K_n}$; (ties broken randomly)
 $Q \leftarrow \frac{(1 - B_m(t))K}{K_m}$; (Q : quantum)
for $n = 1$ **to** N **do**
 $B_n(t) = B_n(t) + Q \frac{K_n}{K}$; (update deficit counters)
end for
 $P(k) \leftarrow m$; (insert source- m in P)
 $B_m(t) \leftarrow 0$;
end for

2) *SAMS Construction:* Note that the optimal frequencies given by (23) depend on the value we choose for $\tilde{c}_n$. However, $\tilde{c}_n$ itself is an artifact of the constructed scheduler. Hence, there is a circular dependency between $\tilde{c}_n$ and the constructed scheduler. Thus, we propose the following fixed point iterative method for the construction of SAMS. As the first step in this construction, we need to find a good initial approximation for $\tilde{c}_n$. For this purpose, we assume that the scheduler is able to transmit source- n packets with a period of $T_n = \frac{s_n}{\tau_n}$ which will ensure a link utilization of τ_n . With this assumption in hand, $\tilde{S}_n$ behaves as $\tilde{S}_n = T_n X_n - S_n$, where X_n is a geometric random variable with parameter u_n . Recalling that a geometrically distributed random variable with parameter u_n has mean $\frac{1}{u_n}$ and variance $\frac{1 - u_n}{u_n^2}$, we have,

$$\mathbb{E}[\tilde{S}_n] = \frac{T_n}{u_n} - s_n, \quad \text{Var}[\tilde{S}_n] = \frac{T_n^2 p_n}{u_n^2} + v_n. \quad (25)$$

In a large-scale status update system, $T_n \gg s_n$ and $T_n^2 \gg v_n$. Hence, we propose to approximate $\tilde{c}_n = \frac{\text{Var}[\tilde{S}_n]}{\mathbb{E}[\tilde{S}_n]^2}$ by p_n which does not depend on τ_n . Therefore, we set our initial approximation of $\tilde{c}_n$ as $\tilde{c}_n^{(0)} = p_n$.

Starting with $\tilde{c}_n^{(0)}$, we run our fixed point iteration algorithm for L iterations. In iteration $\ell, 1 \leq \ell \leq L$, for a given value of $\tilde{c}_n^{(\ell-1)}$, we obtain the source- n frequency $f_n^{*(\ell)}$ according to (23). Then, we choose a $\varepsilon \geq 0$ out of a given set $\mathcal{E}$ of ε values and employ Algorithms 4 and 5 to obtain a pattern $P^{(\ell, \varepsilon)}$ in iteration ℓ whose corresponding weighted AoI $\mathbb{E}[\Delta^{(\ell, \varepsilon)}]$ and $\tilde{c}_n^{(\ell, \varepsilon)}$ values can be calculated based on the procedure described in Appendix IX. We repeat the process for each value ε in the set $\mathcal{E}$ and choose the particular value of ε , denoted by ε' , resulting in a pattern $P^{(\ell)} = P^{(\ell, \varepsilon')}$ with the minimum weighted AoI in the iteration ℓ . Then, the parameter $\tilde{c}_n^{(\ell+1)} = \tilde{c}_n^{(\ell, \varepsilon')}$ is fed as input to iteration $\ell + 1$, and this process is repeated for L iterations. The pattern $P^{(\ell)}$ which generates the lowest weighted AoI among all the iterations $1 \leq \ell \leq L$ is then output by our proposed algorithm, called SAMS (scalable weighted AoI minimizing scheduler), which is a function of the algorithm parameters $\mathcal{E}$ and L . Hence, the proposed method SAMS should not be viewed as a

Algorithm 6 Psuedo Code for SAMS Construction

Input: $\mathcal{E}, L$;
Output: SAMS;
 $\tilde{c}_n^{(0)} = p_n \forall n$
for $\ell = 0$ to $L - 1$ **do**
 $f_n^{*(\ell)} \leftarrow \tilde{c}_n^{(\ell)} \forall n$ (compute optimal frequencies from $\tilde{c}_n^{(\ell)}$)
for $\varepsilon \in \mathcal{E}$ **do**
 $\{K_n^\varepsilon\}_{n=1}^N = \text{ALG 4}(\{f_n^{*(\ell)}\}_{n=1}^N, \varepsilon)$ (obtain the K_n values from Algorithm 4)
 $P^{(\ell, \varepsilon)} = \text{ALG 5}(\{K_n^\varepsilon\}_{n=1}^N)$ (construct the pattern using Algorithm 5)
 $\mathbb{E}[\Delta^{(\ell, \varepsilon)}], \{\tilde{c}_n^{(\ell, \varepsilon)}\}_{n=1}^N \leftarrow P^{(\ell, \varepsilon)}$ (obtain the weighted AoI and the $\tilde{c}_n$ for pattern $P^{(\ell, \varepsilon)}$)
end for
 $\varepsilon' = \arg \min_{\varepsilon} \{\mathbb{E}[\Delta^{(\ell, \varepsilon)}]\}$
 $\tilde{c}_n^{(\ell+1)} = \tilde{c}_n^{(\ell, \varepsilon')}, P^{(\ell)} = P^{(\ell, \varepsilon')}, \mathbb{E}[\Delta^{(\ell)}] = \mathbb{E}[\Delta^{(\ell, \varepsilon')}]$
end for
 $\ell' = \arg \min_{\ell} \{\mathbb{E}[\Delta^{(\ell)}]\}$
SAMS = $P^{(\ell')}$

vector fixed-point iteration but rather a search algorithm along the fixed-point iterations. For a given ε , the computational complexity of one SAMS iteration is $O(NK)$ due to the fact that Algorithm 5 dominates the overall execution time, where K is the size of the pattern produced by SAMS. Algorithm 6 presents the pseudo-code for SAMS construction.

Algorithm 7 Multi-Stage Packet Spreading Algorithm

Input: $N \geq 2, K_n, n = 1, \dots, N, K = \sum_{n=1}^N K_n$;
 $V_0 = [K_1, K_2, \dots, K_N], G_0 = [1, 2, \dots, N], i = 1$
 $G_1, V_1 = \text{GRP}(V_0)$
while $V_i \neq V_{i-1}$ **do**
 $G_{i+1}, V_{i+1} = \text{GRP}(V_i)$
 $i \leftarrow i + 1$
end while
 $P = \text{ALG 5}(V_{i-1})$
for $j = i - 1$ to 1 **do**
 $P = \text{RRA}(P, G_j)$
end for

B. Grouped Packet Spreading Algorithm

In general, Algorithm 5 which is used for packet spreading works well for most input vector of the required source occurrences $\{K_n\}_{n=1}^N$. However, we note that for certain pathological input vectors, its output pattern is not exactly uniform. For example, consider an input vector with 5 sources and the vector of K_n values are $\{8, 1, 1, 1, 1\}$. In this particular scenario, the output pattern of the algorithm would be $P = [1, 1, 1, 1, 1, 1, 1, 1, 2, 3, 4, 5]$. This pattern can be further improved if we can spread the source-1 occurrences uniformly among the other source instances leading to a pattern like $P = [1, 1, 2, 2, 1, 1, 3, 1, 1, 4, 1, 1, 5]$. Therefore, to further improve our packet spreading algorithm we propose the grouped packet spreading algorithm given in Algorithm 7.

In Algorithm 7, $\text{GRP}(V)$ is a function which takes in a vector of the number of source (group) occurrences, groups

together the sources (groups) with the lowest identical number of occurrences and outputs the set of groups G_i along with a vector representing the total number of elements within each group V_i . $\text{RRA}(P, G_j)$ takes in an input set of groups G_j and a pattern P , and allocates the elements present in the l th group of G_j in a round robin fashion at locations in the pattern P reserved for the l th group of G_j .

The general idea behind this algorithm is to group together sources with the same number of occurrences together and treat them as a single source. Once grouped, Algorithm 5 can be used to find the pattern for this grouped source problem. Once this pattern is found, for every instance a particular group appears in pattern, we assign the sources in that particular group in a round robin manner. This allows us to generate more uniformly spread patterns. When the groups are formed, there may be instances where the number of occurrences of two groups are also identical. Therefore, this grouping procedure must be done several times before applying Algorithm 5. This grouped packet spreading algorithm helps further minimize the weighted AoI, but it is at the expense of increased complexity.

As an application of the grouped packet spreading algorithm, let us consider the following input vector of K_n values, $V_0 = [16, 2, 1, 1, 2]$. Using V_0 , the **GRP** function outputs $G_1 = \{[1], [2], [5], [3, 4]\}$ and $V_1 = [16, 2, 2, 2]$. Here, G_1 contains 4 groups where each group contains a vector representing the positions of their corresponding elements within V_0 . Note that in G_1 , only the element 1 in V_0 has been grouped together. Now, if we apply another round of grouping using V_1 , we get $G_2 = \{[1], [2, 3, 4]\}$ and $V_2 = [16, 6]$. At this stage, grouping further would not yield a different vector for V_3 , when compared to V_2 . Therefore, we stop the grouping and apply Algorithm 5 to vector V_2 . This will give us the following pattern $P = [1, 1, 2, 1, 1, 1, 2, 1, 1, 1, 2, 1, 1, 2, 1, 1, 1, 2, 1, 1, 1, 2]$. Now, for each group in G_2 , we allocate the corresponding elements in a round robin fashion. This would not change any elements in P with value 1, but for elements with value 2, it would assign $[2, 3, 4]$ in a round robin manner. This gives us the updated pattern as $P = [1, 1, 2, 1, 1, 1, 3, 1, 1, 1, 4, 1, 1, 2, 1, 1, 1, 3, 1, 1, 1, 4]$. Next, we do another round robin allocation to this new pattern based on the groups in G_1 . This gives the final pattern as $P = [1, 1, 2, 1, 1, 1, 5, 1, 1, 1, 3, 1, 1, 2, 1, 1, 1, 5, 1, 1, 1, 4]$.

VII. ALTERNATIVE AGE-AGNOSTIC SCHEDULING ALGORITHMS

In this section, we describe the modifications we have made on two existing age-agnostic scheduling schemes, namely, the insertion search (IS) and probabilistic GAW (P-GAW) scheduling algorithms, that were first proposed and studied in [29], but for the case when packet errors were absent. We will use them as baseline algorithms (in addition to Eywa [28]) to compare them against NOTS for the two-sources case, and SAMS for general number of sources.

A. Is Algorithm

In this section, we modify and extend the IS algorithm which was introduced in [29] and extend it to accommodate packet drops in the channel. IS is an iterative approach for the construction of a cyclic scheduler where it starts off with

Algorithm 8 Pseudo-Code for IS Algorithm

Input: $\{p_n, s_n, c_n\}$ for $1 \leq n \leq N, I$;
Output: P_{IS}^* ;
 $P^{(N)} = [1, 2, \dots, N]$
 $\mathbb{E}[\Delta^{(N)}] \leftarrow P^{(N)}$ (obtain the weighted AoI $\mathbb{E}[\Delta^{(N)}]$ for pattern $P^{(N)}$)
for $i = N + 1$ **to** I **do**
 for $k = 1$ **to** I **do**
 for $n = 1$ **to** N **do**
 $P_{n,k}^{(i)} = \text{INS}(P^{(i)}, n, k)$
 $\mathbb{E}[\Delta_{n,k}^{(i)}] \leftarrow P_{n,k}^{(i)}$ (obtain the weighted AoI $\mathbb{E}[\Delta_{n,k}^{(i)}]$ for pattern $P_{n,k}^{(i)}$)
 end for
 end for
 $n', k' = \arg \min_{n,k} \mathbb{E}[\Delta_{n,k}^{(i)}]$
 $P^{(i+1)} = P_{n',k'}^{(i)}$
 $\mathbb{E}[\Delta^{(i)}] = \mathbb{E}[\Delta_{n',k'}^{(i)}]$
end for
 $i^* = \arg \min_i \mathbb{E}[\Delta^{(i)}]$
 $P_{IS}^* = P^{(i^*)}$

TABLE I

WORST CASE COMPUTATIONAL COMPLEXITIES OF THE MULTI-SOURCE SCHEDULERS TAKING THE MAXIMUM PERMITTED SCHEDULE SIZE OF IS AS I ($I > K$)

Algorithm	Complexity
IS	$O(N^2 I^3)$
SAMS	$O(NK)$
Eywa	$O(N^2 K^2)$

an RR pattern and iteratively expands this pattern. Suppose the current pattern and pattern size in the i th iteration of the IS algorithm are $P^{(i)}$ and $K^{(i)}$, respectively. Then, in each iteration, the IS algorithm selects a source from among the N sources and a location from among the $K^{(i)}$ possible locations, where a new scheduling instance of the selected source can be inserted into the current pattern $P^{(i)}$, to construct a total of $NK^{(i)}$ patterns. The pattern which results in the lowest weighted AoI is then selected as the starting pattern for the next iteration, $P^{(i+1)}$. This procedure is repeated for I iterations and the pattern that resulted in the lowest weighted AoI is selected from the patterns $P^{(i)}$ for $i = 1$ to I . One difference when compared to the IS algorithm presented in [29], is that we do not terminate the algorithm when the pattern in the current iteration has a higher weighted AoI than the pattern in the previous iteration. Instead, we continue for I iterations and find the best possible pattern. The computational complexity of the IS algorithm is $O(N^2 I^3)$ and its pseudo-code is presented in Algorithm 8, where $\text{INS}(P, n, k)$ is a function which takes in a pattern P and inserts a new source- n scheduling instance after the k th element in the pattern P and outputs this new pattern. Table I compares the worst case computational complexities of IS along with the other main multi-source scheduling schemes ($N > 2$) considered in this work.

B. Probabilistic Scheduler

The probabilistic generate-at-will (P-GAW) model was introduced in [29] for scheduling N sources in the absence of packet drops, where at every scheduling instance, source- n is selected based on a probability η_n , such that $\sum_{n=1}^N \eta_n = 1$. In this section, we show how we can extend the results in [29] to find the weighted AoI of a P-GAW model in the presence of packet drops. For this purpose, we treat the system of N sources with packet drops as a system with $N + 1$ sources without packet drops by considering a phantom source which incorporates all the service times that resulted in a packet drop.

Let source- $(N + 1)$ represent the phantom source and let η'_n represent the probability that source- n packet transmits a packet successfully (without drop). Therefore, η'_n corresponds to the event where source- n was chosen for transmission and its packet was successfully transmitted. This gives us $\eta'_n = \eta_n(1 - p_n)$. Whenever a scheduled transmission fails, we assume that transmission was occupied by the phantom source. Therefore, source- $(N + 1)$ transmission corresponds to events where the actual scheduled transmission has failed. This gives us $\eta'_{N+1} = \sum_{n=1}^N \eta_n p_n$. The first and second moments of the service time of the phantom source are given as follows,

$$s_{N+1} = \frac{1}{\eta'_{N+1}} \sum_{n=1}^N \eta_n p_n s_n, \quad q_{N+1} = \frac{1}{\eta'_{N+1}} \sum_{n=1}^N \eta_n p_n q_n. \quad (26)$$

Therefore, the weighted AoI of the actual N -source system in the presence of packet drops, scheduled according to a P-GAW scheduler with selection probabilities η_n , is equivalent to the weighted AoI of the $(N + 1)$ -source system in the absence of packet drops, scheduled according to a P-GAW scheduler with selection probabilities η'_n . Hence, the AoI analysis of the P-GAW model presented in [29] can directly be applied to this phantom-source expanded $(N + 1)$ -source system.

VIII. NUMERICAL RESULTS

In this section, we evaluate the performance of our constructed schedulers and benchmark them against the existing multi-source schedulers whenever possible. We present three variations of the proposed SAMS algorithm that are implemented in our numerical experiments. In SAMS-1, the algorithm parameters are selected as $\mathcal{E} = \{0\}$, $L = 1$ whereas SAMS-2 employs a wider set $\mathcal{E} = \{0 : 0.2 : 2\}$ for $\mathcal{E}$, but does not perform fixed-point iterations. Finally, SAMS-3 uses the same set $\mathcal{E}$, but employs fixed-point iterations with $L = 3$. In all numerical experiments, SAMS uses the ungrouped version of the packet spreading algorithm unless specifically stated otherwise. We first consider the case when $N = 2$ and compare the performance of NOTS with the optimal P-GAW scheduler denoted by P-GAW*, RR (round robin) policy, IS algorithm and SAMS-3 for two sources with exponentially distributed source service times. Here, NOTS is constructed using $\alpha = 50$, whereas P-GAW* is implemented by finding the optimal η_n probabilities through exhaustive search.

In our first experiment, we fix the parameters of source-2 and vary either the packet drop probability or the mean service time of source-1. As seen in Fig. 4, NOTS achieves significantly lower weighted AoI than the P-GAW* and the RR policy. This is to be expected since the weighted AoI of the P-GAW model is simply an expectation of all deterministic

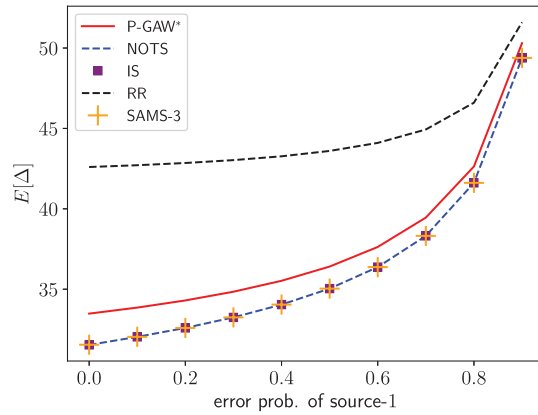
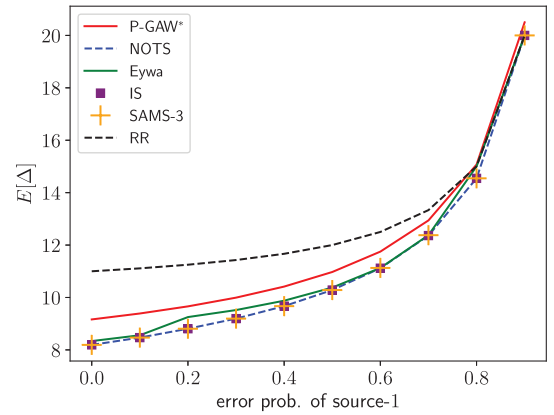
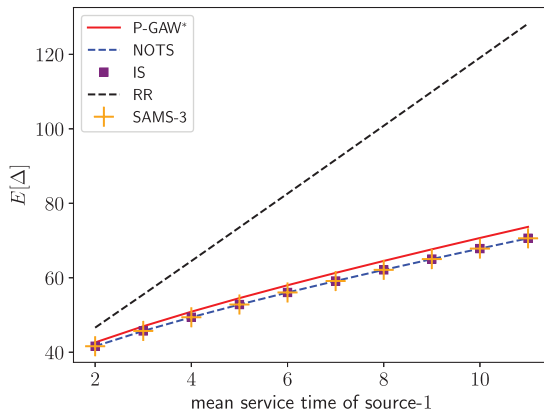
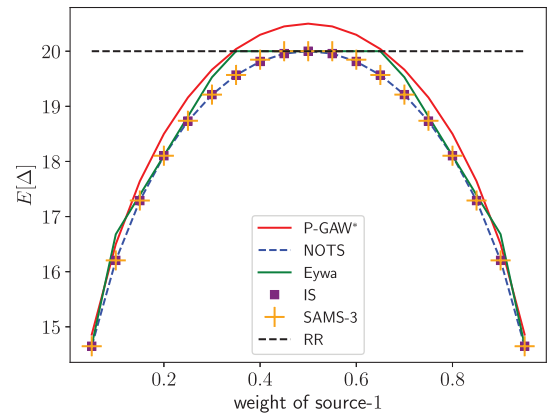
(a) $s_1 = 2$ (a) $s_1 = s_2 = 1, w_1 = 0.5, w_2 = 0.5, p_2 = 0.9$ (b) $p_1 = 0.8$ (b) $s_1 = s_2 = 1, p_1 = 0.9, p_2 = 0.9$

Fig. 4. Weighted AoI of a two source system depicted as a function of (a) packet drop probability, (b) mean service time, of source-1, for exponential service times ($s_2 = 3, w_1 = 0.2, w_2 = 0.8, p_2 = 0.9$).

Fig. 6. Variation of the weighted AoI of a two source system with respect to (a) packet drop probability, (b) weight, of source-1, for deterministic service times.

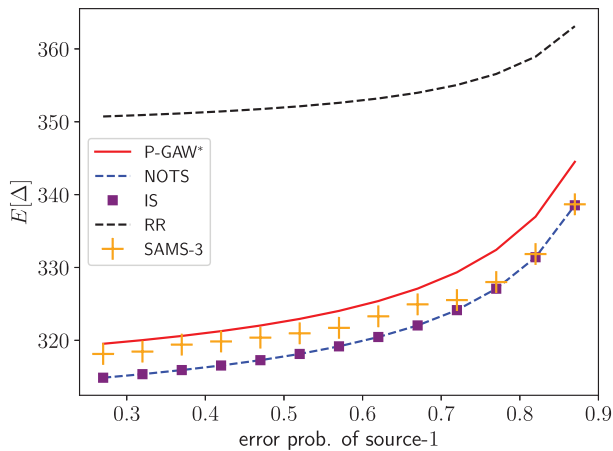
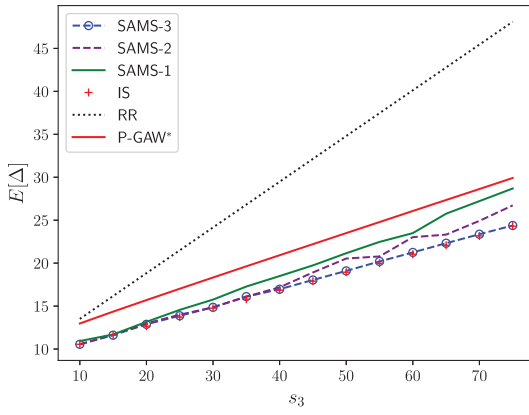


Fig. 5. Variation of weighted AoI with the packet drop probability of source-1 for a scenario with highly variable service times ($s_1 = 25, s_2 = 24, c_1 = 2, c_2 = 15, p_2 = 0.81, w_1 = 0.04, w_2 = 0.96$).

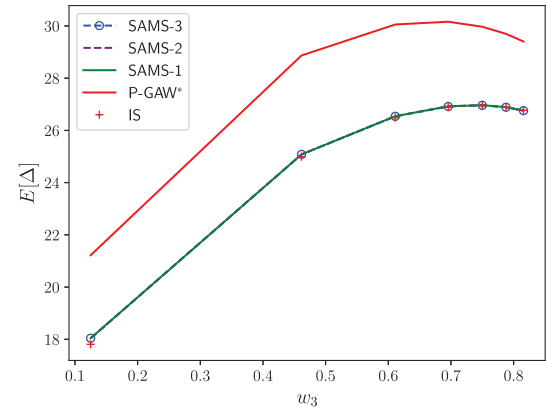
cyclic schedules, whereas NOTS is the closest to the best cyclic schedule and is constructed by considering policies which are better than a simple RR policy. Both IS and SAMS-3 are as good as NOTS for this particular case. However, this is not observed in all scenarios. Fig. 5 shows one such scenario

for which there is high variability in source service times, i.e., relatively large c_n values. In this case, even though SAMS-3 deviates from NOTS, IS still generates a near-optimal schedule. One of the key reasons for the performance degradation of SAMS-3 in this scenario is the initial approximation of $\tilde{c}_n = p_n$ holding well for deterministic (or deterministic-like) service times with relatively low c_n values.

Next, we compare the performance of the above schedulers for identical deterministic source service times (i.e., $s_1 = s_2$ and $c_1 = c_2 = 0$). Since this experimental setup coincides with that of Eywa, we have included Eywa as a benchmark policy for this particular experiment. We fix the parameters of source-2, and in one instance, we vary the packet drop probability of source-1, and in the other instance, we vary the weight associated with the two sources. Fig. 6 shows that NOTS outperforms Eywa whereas SAMS-3 closely follows NOTS. [29] shows that IS is optimal for two sources in the absence of packet errors, and our experimental results show that IS performs very close to the optimum solution when $N = 2$. However, the drawback of IS is its higher computational complexity and the lack of provable guarantees for its performance in the presence of packet errors when $N > 2$.



(a) $s_1 = 5$, $s_2 = 2.5$, $p_1 = p_2 = p_3 = 0$,
 $w_1 = 5w_2 = 25w_3$



(b) $s_1 = 10$, $s_2 = s_3 = 1$, $p_1 = 0.1$, $p_2 = 0.5$,
 $p_3 = 0.95$, $w_2 = 2.5w_1$

Fig. 7. Variation of the weighted AoI of a three source system with respect to (a) packet drop probability, (b) weight, of source-3, for deterministic service times.

Now, we consider the case $N > 2$, and compare the performance of the three variants of the SAMS algorithm. We first study a small scale status update system with $N = 3$ with deterministic service times. In this experiment, we fix the parameters of the first two sources and vary the mean and the weight of the third source. As a benchmark policy, we use the IS algorithm with the choice of $I = 75$ since the above empirical evidence suggests that it is close to optimal for $N = 2$. Fig. 7a depicts the variation of the weighted AoI with the mean service time of the third source s_3 for a highly heterogeneous scenario with no packet drops. As illustrated, the RR policy performs very poorly especially when heterogeneity increases, i.e., s_3 is increased. The performance gap between P-GAW* and IS is also substantial for relatively large s_3 values. The results obtained with the three variations of SAMS reveal that: (i) exhaustive search over a number of ε values is advantageous compared to using a single value of this parameter (observe that SAMS-2 outperforms SAMS-1); (ii) with the addition of a few fixed-point iterations, the weighted AoI performance can further be improved (observe that SAMS-3 outperforms SAMS-2). Overall, for this example, the performance of SAMS-3 has been very close to that of IS. Fig. 7b shows the variation of the weighted AoI with the weight allocated for the third source w_3 for a system with non-zero packet drop probabilities. The ratio of the weights between first two sources is fixed as we vary w_3 in this example. In this example, all three variations of the SAMS algorithm closely follow the IS algorithm while exhibiting a significant performance gain over P-GAW*. The RR policy is purposefully avoided in this example, since it is inherently worse compared to all the other policies. Note that the computational complexity of IS is $O(N^2I^3)$ which is far higher than the worst case complexity of $O(NK)$ for SAMS (assuming the produced pattern size K in SAMS is less than I), the latter being applicable to very large-scale scenarios as well. In this regard, SAMS-3 performance comes very close to IS for deterministic service times, despite the substantial gap in their computational complexities.

To compare how grouping affects the packet spreading algorithm, we evaluate the performance of the SAMS-3 algorithm with its grouped counterpart, SAMS-3G. SAMS-3G uses the

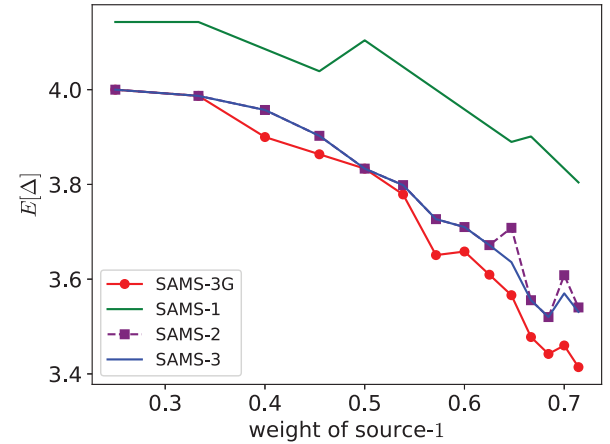


Fig. 8. Variation of the weighted AoI with respect to the weight of source-1 for i.i.d exponentially distributed service times and zero packet error probabilities ($w_2 = w_3 = 2w_4 = 2w_4$, $s_i = 1$, $\forall i$).

same parameters for L and $\mathcal{E}$ that were used in SAMS-3, with the only exception being the packet spreading algorithm. As the packet spreading algorithm in SAMS-3G, we use Algorithm 7 as opposed to Algorithm 5 employed in SAMS-3. In this experiment, we consider a system with five sources with i.i.d. exponential services times in the absence of packets errors but with different weights allocated to each source. Fig. 8 shows the variation of the weighted AoI with the weight of source-1. As shown in Fig. 8, there is a clear performance gain when using the grouped packet spreading algorithm for this small-scale status update system. However, we avoid using the grouped packet spreading algorithm for large-scale status update systems due to the added complexity introduced by grouping.

Next, to evaluate the performance of the SAMS algorithm for large-scale status update systems. We compare the performance of SAMS against the Eywa framework which is designed only for systems with deterministic and identical service times. We have intentionally left out the P-GAW* scheduler for this large-scale system, since the optimal proba-

TABLE II
VARIATION OF THE PATTERN SIZE AND THE COMPUTATION TIME IN SECONDS OF THE SAMS-3 ALGORITHM FOR VARIOUS REPRESENTATIVE MASSIVE SCALE SCENARIOS

N	MS1		MS2		MS3		MS4	
	K	Comp. time	K	Comp. time	K	Comp. time	K	Comp. time
128	874	32	1198	71	1041	83	885	49
256	1748	121	2384	282	2304	271	1523	124
512	3283	517	4768	1244	4766	1282	3496	589
1024	6984	2879	9534	7459	8297	8833	7047	3475

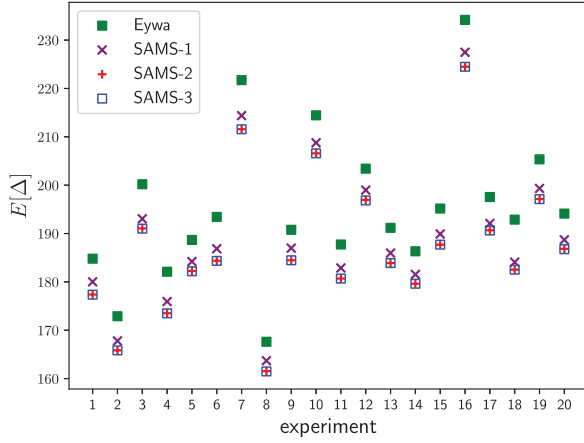
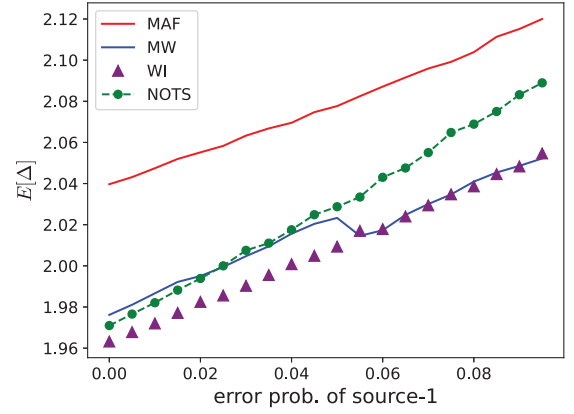


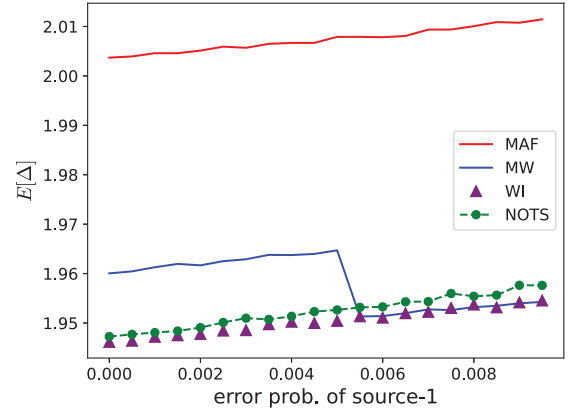
Fig. 9. Weighted AoI for $N = 100$ with deterministic and identical service times for twenty different experiments.

bilities of the probabilistic scheduler is found via an exhaustive search, and therefore would be infeasible in these large-scale scenarios. To bring our model to the same domain as Eywa, we set the service times of all sources to 1 with $c_n = 0$, $\forall n$. We take $N = 100$ sources, and randomly sample 20 weight vectors and probability of error vectors, and evaluate the weighted AoI of the two schemes, namely SAMS (with its three variations) and Eywa, in each case. As illustrated in Fig. 9, in each of the experiments, SAMS outperforms Eywa by a significant margin. This performance gain is achieved without any sacrifice in the computational complexity. In fact, the computational complexity of Eywa is $O(N^2 K^2)$ where K in here is the maximum pattern size generated by Eywa with $K < I$. Moreover, in this example, SAMS-2 and SAMS-3 performances are very close. Actually, the $\tilde{c}_n$ values obtained by the SAMS-2 scheduler are already very close to p_n , and therefore, there is not much need for performing fixed-point iterations specific to SAMS-3. However, SAMS-2 has a considerable performance gain over SAMS-1, indicating that even though fixed point iterations are not much useful in these large-scale systems, the local search mechanisms (finding the optimal ε) in SAMS-2 is advantageous. Finally, the computational complexity of Eywa is much higher than SAMS and this is the sole reason for limiting this particular example only to 20 experiments and 100 sources.

In fact, it is quite possible to obtain well-performing schedulers for up to thousand information sources with SAMS, thanks to its low computational complexity. Now, we present the results for systems representative of massive IoT networks, i.e., $N > 100$. For these experiments, we evaluate how the



(a) $p_1 \leq 0.1$, $p_2 = 0.05$



(b) $p_1 \leq 0.01$, $p_2 = 0.005$

Fig. 10. Comparison of NOTS with age-aware policies for deterministic and identical service times ($w_1 = \frac{5}{6}$, $w_2 = \frac{1}{6}$).

pattern size and the pattern computation time vary with the number of sources for massive scale (MS) scenarios named MS1, MS2, MS3 and MS4. In MS1, we consider identical deterministic service times $s_n = 1$, zero packet drop probabilities $p_n = 0$ and linearly increasing source weights $w_n = nw_1$. MS2 is similar to MS1 with the exception that the packet drop probabilities of the sources are evenly spaced between 0 to 0.5 as $p_n = \frac{1}{2n}$. MS3 follows MS1 but with non-identical deterministic service times where $s_n = n \bmod 4 + 1$. Finally, MS4 uses identical exponentially distributed service times $s_n = 1$, $c_n = 1$, with the other parameters chosen the same as MS1. Each of these experiments is run on a Python 3 Google Compute Engine with a 2.30GHz CPU and 12GB RAM. Table II summarizes the results for the four studied

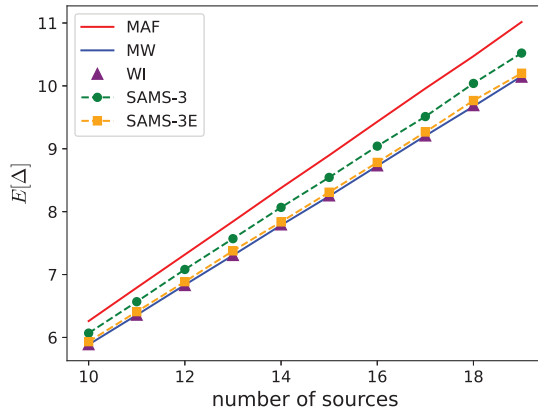
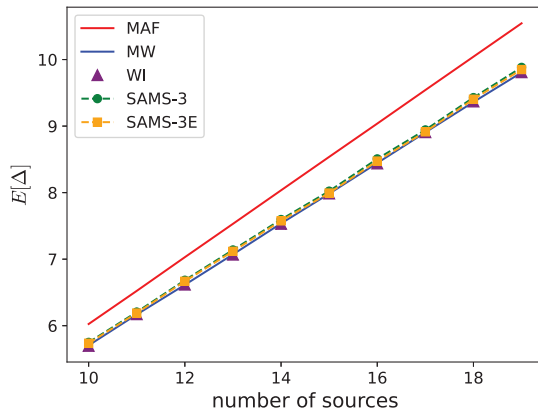
(a) $p_n = \frac{(n-1)}{100N} \leq 0.1$ (b) $p_n = \frac{(n-1)}{100N} \leq 0.01$

Fig. 11. Comparison of SAMS-3 with age-aware policies for deterministic and identical service times with $w_n = \frac{2(N-n)}{N(N+1)}$.

scenarios where we see that, even with ordinary computing resources, one can obtain a schedule for $N = 1024$ sources in a couple of hours. This indicates that constructing schedules for thousands of sources is well within the reach of our proposed algorithms. It is worth noting that, despite the high computational execution times observed for constructing the schedules for large number of sources, once a schedule has been constructed, the runtime complexity of the schedule is extremely low, i.e., $O(1)$, due to the simplicity of running a cyclic scheduler.

Finally, we compare our cyclic scheduling schemes against several existing age-aware policies. For this purpose, we use the maximum age first (MAF), max-weight (MW) and the Whittle index (WI) policies introduced in [22] which are the best-known multi-source scheduling policies for identical deterministic service times and they are known to be optimal for symmetric systems, i.e., identical packet drop probabilities and weights. These three existing policies are designed for the discrete-time setting where each service time is assumed to be 1. Therefore, we evaluate our proposed schedulers under the setting where $s_n = 1$ and $c_n = 0$ (identical deterministic service times). Moreover, for a fair comparison, we consider the regime of low error probabilities across all sources in these experiments since the cyclic schemes were designed

specifically with the assumption that there is no error feedback. This is a reasonable assumption since in most practical communication scenarios, the packet error probabilities are reasonably low. Fig. 10 illustrates the variation of the weighted AoI as we vary the packet drop probability of source-1 for a two-source scenario with fixed source weights and fixed drop probability of source-2. As illustrated in Fig. 10a, NOTS exhibits reasonable performance compared to the WI policy and even outperforms the MW policy for very low error probabilities. When the error probabilities are lower as in Fig. 10b, NOTS follows the WI policy even more closely.

Next, we compare the performance of SAMS-3 against the same three age-aware policies when we vary the number of sources. In this experiment, we also consider another variant of SAMS-3 termed as SAMS-3E which is assumed to have feedback on packet errors and in the event of a packet drop, it retransmits from the same source by taking a new sample, and moves to the next source in the schedule only after the current source has a successful transmission. As seen in Fig. 11a, SAMS-3 has a reasonable performance gap when compared to the WI and MW policies while SAMS-3E closely follows the age-aware policies. This is expected, since with the added error feedback, we can make better scheduling decisions (for high error probabilities SAMS-3E may not be a good candidate since the AoI of other sources will increase drastically while waiting for a successful transmission of the current source). As seen in Fig. 11b, when the packet error probabilities are reduced further, SAMS-3 also closely follows the age-aware policies.

IX. CONCLUSION

In this work, we have first introduced two novel analytical frameworks to compute the weighted AoI of a multi-source status update system with N sources employing a cyclic scheduler, on the basis of which we have proposed an algorithm called SAMS (scalable weighted AoI Minimization Scheduler) for constructing cyclic schedules for up to thousand sources, which can be employed in massive-size networks such as large-scale IoT networks. We have obtained a theoretical lower bound achievable by cyclic scheduling for $N = 2$, and we presented an algorithm, called NOTS, that can be close to this bound as desired. The lower bound presented in this work enables us to evaluate the optimality of a number of heuristic age-agnostic scheduling algorithms for the specific case of $N = 2$ sources. For small-scale status update systems, we have shown that SAMS performs very close to the optimum solution for $N = 2$, and also very close to the benchmark IS (insertion search) algorithm for deterministic service times for $N = 3$. For moderately large-scale status update systems, i.e., $N = 100$, SAMS is shown to perform better than the best-known age-agnostic scheduler existing in the literature that can be used for $N = 100$, with a lesser computational complexity. We have also shown that cyclic schedule construction for massive-scale status update systems with up to thousand sources is possible with the proposed SAMS algorithm, while no other existing age-agnostic scheduling algorithm appears to be feasible in this massive-connectivity regime.

Future extensions of this work may involve finding good performance bounds on the proposed algorithms, improved packet spreading algorithms for random service times,

modeling update systems with non-located server and monitor, non-negligible polling delays, sources with random arrivals (RA) of packets as opposed to generate-at-will (GAW) systems, sources with duty cycle constraints, and so on. Applications of cyclic scheduling to systems with inherent periodic characteristics such as a server taking vacations based on a periodic schedule, is another interesting line of research.

APPENDIX A PROOF OF THEOREM 1

Let $A_{i,j,k}$ represent the area of the k th AoI cycle starting at the i th scheduling instance of source- n and ending in the j th scheduling instance of source- n . Let $T_{i,j,k}$ represent the cycle time of the same AoI cycle. Let $N_{i,j}$ be the number of AoI cycles which start at the i th scheduling instance of source- n and end in the j th scheduling instance of source- n within a time horizon of T . Let N_T be the total number of AoI cycles within a time horizon of T . Therefore, $N_T = \sum_{i,j} N_{i,j}$. Then, the time averaged AoI of source- n , denoted by $\bar{\Delta}_n$ which is also equal to $\mathbb{E}[\Delta_n]$, can be found as follows,

$$\bar{\Delta}_n = \limsup_{T \rightarrow \infty} \frac{\int_0^T \Delta_n(t) dt}{T} \quad (27)$$

$$= \limsup_{N_T \rightarrow \infty} \frac{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \sum_{k=1}^{N_{i,j}} A_{i,j,k}}{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \sum_{k=1}^{N_{i,j}} T_{i,j,k}} \quad (28)$$

$$= \lim_{N_T \rightarrow \infty} \frac{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \frac{N_{i,j}}{N_T} \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} A_{i,j,k}}{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \frac{N_{i,j}}{N_T} \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} T_{i,j,k}} \quad (29)$$

$$= \frac{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \lim_{N_T \rightarrow \infty} \frac{N_{i,j}}{N_T} \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} A_{i,j,k}}{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \lim_{N_T \rightarrow \infty} \frac{N_{i,j}}{N_T} \frac{1}{N_{i,j}} \sum_{k=1}^{N_{i,j}} T_{i,j,k}} \quad (30)$$

$$= \frac{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \pi_{i,j} \mathbb{E}[A_{i,j}]}{\sum_{i=1}^{\alpha_n} \sum_{j=1}^{\alpha_n} \pi_{i,j} \mathbb{E}[T_{i,j}]} \quad (31)$$

In this derivation, (28) holds almost surely (the surplus area not captured by an AoI drop is finite with probability one) and it reduces to (29) if the limit exists. If the limit exists in (29), it will reduce to (30) if the limits of the numerator and the denominator exist and are finite. The limits of the numerator and denominator in (30) reduce to that of (31) from the law of large numbers and the fact that the MC is ergodic. From (7) and (8), we can conclude that the numerator and denominator of (31) are finite and hence assert the correctness of the logical flow from (28) to (30).

APPENDIX B MGF APPROACH

Let P be the pattern for the given cyclic schedule for N information sources. We now present the MGF method to derive the average AoI of source- n through (10) which requires the quantities $\tilde{s}_n$ and $\tilde{q}_n$ for which purpose we define $P_{n,k}$, $k = 0, 1, \dots, \alpha_n - 1$, as the sub-pattern between the k th and $(k+1)$ th (modulo α_n) occurrences of source- n within P , excluding the end points. In what follows, the k th and $(k+\alpha_n)$ th occurrences of source- n refer to the same occurrence since the patterns repeat themselves. Let $\alpha_{n,k,m}$ denote the number of times source- m occurs in the sub-pattern $P_{n,k}$. Let us denote by $H_{n,k}$ the random variable corresponding to the sum of the

service times of all sources contained within the sub-pattern $P_{n,k}$ with mean $s_{n,k} = \mathbb{E}[H_{n,k}]$, second moment $q_{n,k} = \mathbb{E}[H_{n,k}^2]$, and MGF $G_{n,k}(s) = \mathbb{E}[e^{sH_{n,k}}]$. Since all the service times are independent, it is easy to write,

$$G_{n,k}(s) = \prod_{m \neq n} G_m(s)^{\alpha_{n,k,m}}, \quad (32)$$

$$s_{n,k} = \sum_{m \neq n} s_m \alpha_{n,k,m}, \quad (33)$$

$$q_{n,k} = \sum_{m \neq n} v_m \alpha_{n,k,m} - s_{n,k}^2. \quad (34)$$

Now, let $\tilde{H}_{n,k}$ denote the time between two consecutive successful transmissions of source- n (excluding the end points) given that the first successful transmission occurs at the k th occurrence of source- n in pattern P with mean $\tilde{s}_{n,k} = \mathbb{E}[\tilde{H}_{n,k}]$, second moment $\tilde{q}_{n,k} = \mathbb{E}[\tilde{H}_{n,k}^2]$, variance $\tilde{v}_{n,k} = \tilde{q}_{n,k} - \tilde{s}_{n,k}^2$ and MGF $\tilde{G}_{n,k}(s) = \mathbb{E}[e^{s\tilde{H}_{n,k}}]$. A successful source- n packet belongs to a transmission at occurrence- k , $k = 0, 1, \dots, \alpha_n$, uniformly likely. Therefore, the following identities hold,

$$\tilde{G}_n(s) = \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \tilde{G}_{n,k-1}(s), \quad (35)$$

$$\tilde{s}_n = \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \tilde{s}_{n,k}, \quad (36)$$

$$\tilde{q}_n = \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \tilde{q}_{n,k}. \quad (37)$$

Let us first compute the quantity $\tilde{s}_{n,k}$. Using the total law of expectation, we have

$$\begin{aligned} \tilde{s}_{n,k}/u_n &= s_{n,k} + p_n(s_n + s_{n,k} + s_{n,k+1}) + \dots \\ &+ p_n^{\alpha_n-1}((\alpha_n - 1)s_n + s_{n,k} + s_{n,k+1} + \dots + s_{n,k+\alpha_n-1}) \\ &+ p_n^{\alpha_n}(\alpha_n s_n + 2 s_{n,k} + s_{n,k+1} + \dots + s_{n,k+\alpha_n-1}) \\ &+ p_n^{\alpha_n+1}((\alpha_n + 1)s_n + 2 s_{n,k} + 2 s_{n,k+1} \\ &\quad + s_{n,k+2} + \dots + s_{n,k+\alpha_n-1}) + \dots \\ &+ p_n^{2\alpha_n-1}((2\alpha_n - 1)s_n + 2 s_{n,k} \\ &\quad + 2 s_{n,k+1} + \dots + 2 s_{n,k+\alpha_n-1}) \\ &+ p_n^{2\alpha_n}(2\alpha_n s_n + 3 s_{n,k} \\ &\quad + 2 s_{n,k+1} + \dots + 2 s_{n,k+\alpha_n-1}) + \dots \end{aligned} \quad (38)$$

In the above identity, we collect the terms belonging to s_n and $s_{n,k+j-1}$, $1 \leq j \leq \alpha_n$, to show the following,

$$\tilde{s}_{n,k} = \frac{s_n p_n}{u_n} + \frac{u_n \left(\sum_{j=1}^{\alpha_n} p_n^{j-1} s_{n,k+j-1} \right)}{(1 - p_n^{\alpha_n})^2} \sum_{j=1}^{\alpha_n} p_n^{j-1} \quad (39)$$

$$= \frac{s_n p_n}{u_n} + \frac{\sum_{j=1}^{\alpha_n} p_n^{j-1} s_{n,k+j-1}}{1 - p_n^{\alpha_n}}. \quad (40)$$

From the above expression and also using (36), we obtain the following closed-form expression for $\tilde{s}_n$,

$$\tilde{s}_n = \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \left(\frac{p_n s_n}{u_n} + \frac{\sum_{j=1}^{\alpha_n} p_n^{j-1} s_{n,k+j-1}}{1 - p_n^{\alpha_n}} \right) \quad (41)$$

$$= \frac{1}{u_n} \left(p_n s_n + \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} s_{n,k-1} \right). \quad (42)$$

For the computation of $\tilde{q}_n$, we use the law of total expectation again to write the MGF $\tilde{G}_{n,k}(s)$ as,

$$\begin{aligned} \frac{\tilde{G}_{n,k}(s)}{u_n} &= G_{n,k}(s) + p_n G_n(s) G_{n,k}(s) G_{n,k+1}(s) \\ &\quad + p_n^2 G_n(s)^2 G_{n,k}(s) G_{n,k+1}(s) G_{n,k+2}(s) + \dots \\ &\quad + p_n^{\alpha_n-1} G_n(s)^{\alpha_n-1} \prod_{m=1}^{\alpha_n} G_{n,m-1}(s) \\ &\quad + p_n^{\alpha_n} G_n(s)^{\alpha_n} G_{n,k}(s) \prod_{m=1}^{\alpha_n} G_{n,m-1}(s) \\ &\quad + \left(p_n^{\alpha_n+1} G_n(s)^{\alpha_n+1} G_{n,k}(s) \right. \\ &\quad \quad \left. \times G_{n,k+1}(s) \prod_{m=1}^{\alpha_n} G_{n,m-1}(s) + \dots \right) \quad (43) \\ &= \underbrace{\sum_{j=1}^{\alpha_n} p_n^{j-1} G_n(s)^{j-1} \prod_{l=1}^j G_{n,k+l-1}(s)}_{X_{n,k}(s)} \\ &\quad + p_n^{\alpha_n} G_n(s)^{\alpha_n} \prod_{m=1}^{\alpha_n} G_{n,m-1}(s) \underbrace{X_{n,k}(s)}_{\Pi_n(s)} \\ &\quad + p_n^{2\alpha_n} G_n(s)^{2\alpha_n} \Pi_n(s)^2 X_{n,k}(s) + \dots \quad (44) \end{aligned}$$

We note that the above equation can be simplified as follows,

$$\tilde{G}_{n,k}(s) = \frac{u_n X_{n,k}(s)}{1 - p_n^{\alpha_n} G_n(s)^{\alpha_n} \Pi_n(s)}. \quad (45)$$

Ignoring the higher order terms (since they will not have any effect on $\tilde{G}_{n,k}''(0)$), we write,

$$G_n(s) = 1 + s_n s + \frac{q_n}{2} s^2 + O(s^3), \quad (46)$$

$$G_{n,k}(s) = 1 + s_{n,k} s + \frac{q_{n,k}}{2} s^2 + O(s^3), \quad (47)$$

from which we can write,

$$\tilde{G}_{n,k}(s) = \frac{(1 - p_n^{\alpha_n}) + a_{n,k} s + b_{n,k} s^2 + O(s^3)}{(1 - p_n^{\alpha_n}) + c_{n,k} s + d_{n,k} s^2 + O(s^3)}, \quad (48)$$

where the coefficients $a_{n,k}$, $b_{n,k}$, $c_{n,k}$ and $d_{n,k}$ can be found numerically, e.g., by applying the convolution operator for the product of MGFs. Then, by differentiating (48) twice and evaluating at $s = 0$, the $O(s^3)$ terms will vanish, yielding $\tilde{q}_n$ as follows,

$$\tilde{q}_n = \mathbb{E}[\tilde{s}_n^2] = \tilde{G}_{n,k}''(0) = \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \tilde{G}_{n,k-1}''(0) \quad (49)$$

$$= \frac{1}{\alpha_n} \sum_{k=1}^{\alpha_n} \left(\frac{2c_{n,k}(c_{n,k} - a_{n,k})}{(1 - p_n^{\alpha_n})^2} + \frac{2(b_{n,k} - d_{n,k})}{(1 - p_n^{\alpha_n})} \right). \quad (50)$$

Substituting the values of $\tilde{s}_n$ and $\tilde{q}_n$ obtained in (42) and (50), respectively, into (10), gives the average AoI of source- n .

APPENDIX C

PROOF OF THEOREM 2 AND COROLLARY 1

To obtain the mean AoI of source-1, we need to first compute the terms $\sum_{i,j} \pi_{i,j} \bar{z}_{i,j}$ and $\sum_{i,j} \pi_{i,j} \bar{z}_{i,j}^2$. By exploiting symmetry in stationary distribution of the MC, we can obtain the first summations as follows,

$$\bar{z}_{i,j} = \begin{cases} (j-i)s_1 + \sum_{k=i}^{j-1} \gamma_k s_2, & \text{if } j > i, \\ (\alpha_1 - i + j)s_1 + \sum_{k=i}^{j+\alpha_1-1} \gamma_k s_2, & \text{if } i \geq j, \end{cases} \quad (51)$$

and

$$\begin{aligned} \sum_{i=1}^{\alpha_1} \sum_{j=1}^{\alpha_1} \pi_{i,j} \bar{z}_{i,j} &= \sum_{i=1}^{\alpha_1} \left(\sum_{j=1}^{\alpha_1-i} \pi_{j,j+i} \bar{z}_{j,j+i} + \sum_{j=\alpha_1-i+1}^{\alpha_1} \pi_{j,j+i-\alpha_1} \bar{z}_{j,j+i-\alpha_1} \right) \quad (52) \\ &= \sum_{i=1}^{\alpha_1} \frac{(1-p_1)p_1^{i-1}}{\alpha_1(1-p_1^{\alpha_1})} \left(\sum_{j=1}^{\alpha_1-i} \bar{z}_{j,j+i} + \sum_{j=\alpha_1-i+1}^{\alpha_1} \bar{z}_{j,j+i-\alpha_1} \right) \quad (53) \end{aligned}$$

$$\begin{aligned} &= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\sum_{j=1}^{\alpha_1-i} \left(i s_1 + s_2 \sum_{k=j}^{j+i-1} \gamma_k \right) \right. \\ &\quad \left. + \sum_{j=\alpha_1-i+1}^{\alpha_1} \left(i s_1 + s_2 \sum_{k=j}^{j+i-1} \gamma_k \right) \right) \quad (54) \\ &= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\alpha_1 i s_1 + s_2 \sum_{j=1}^{\alpha_1} \sum_{k=j}^{j+i-1} \gamma_k \right) \quad (55) \\ &= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\alpha_1 i s_1 + s_2 \sum_{k=1}^{\alpha_1} i r_k \right) \quad (56) \end{aligned}$$

$$\begin{aligned} &= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} (\alpha_1 i s_1 + \alpha_2 i s_2) \quad (57) \\ &= s \frac{(1-p_1)}{(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} i p_1^{i-1}. \quad (58) \end{aligned}$$

Similarly, we can obtain the second summation as follows,

$$\begin{aligned} \bar{z}_{i,j}^2 &= \text{Var}(Z_{i,j}) + \bar{z}_{i,j}^2, \\ &= \begin{cases} (j-i)v_1 + \sum_{k=i}^{j-1} \gamma_k v_2 + \bar{z}_{i,j}^2, & \text{if } j > i, \\ (\alpha_1 - i + j)v_1 + \sum_{k=i}^{j+\alpha_1-1} \gamma_k v_2 + \bar{z}_{i,j}^2, & \text{if } i \geq j, \end{cases} \quad (59) \end{aligned}$$

and

$$\begin{aligned} \sum_{i=1}^{\alpha_1} \sum_{j=1}^{\alpha_1} \pi_{i,j} \bar{z}_{i,j}^2 &= \sum_{i=1}^{\alpha_1} \left(\sum_{j=1}^{\alpha_1-i} \pi_{j,j+i} \bar{z}_{j,j+i}^2 + \sum_{j=\alpha_1-i+1}^{\alpha_1} \pi_{j,j+i-\alpha_1} \bar{z}_{j,j+i-\alpha_1}^2 \right) \quad (60) \end{aligned}$$

$$= \sum_{i=1}^{\alpha_1} \frac{(1-p_1)p_1^{i-1}}{\alpha_1(1-p_1^{\alpha_1})} \left(\sum_{j=1}^{\alpha_1-i} \tilde{z}_{j,j+i}^2 + \sum_{j=\alpha_1-i+1}^{\alpha_1} \tilde{z}_{j,j+i-\alpha_1}^2 \right) \quad (61)$$

$$= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\sum_{j=1}^{\alpha_1-i} \left(is_1 + s_2 \sum_{k=j}^{j+i-1} \gamma_k \right)^2 + \sum_{j=\alpha_1-i+1}^{\alpha_1} \left(is_1 + s_2 \sum_{k=j}^{j+i-1} \gamma_k \right)^2 \right) \quad (62)$$

$$= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\alpha_1 i^2 s_1^2 + 2is_1 s_2 \sum_{j=1}^{\alpha_1} \sum_{k=j}^{j+i-1} \gamma_k + s_2^2 \sum_{j=1}^{\alpha_1} \left(\sum_{k=j}^{j+i-1} \gamma_k \right)^2 \right) \quad (63)$$

$$= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\alpha_1 i^2 s_1^2 + 2i^2 \alpha_2 s_1 s_2 + s_2^2 \tilde{r}(i) \right) \quad (64)$$

$$= \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} \left(\alpha_1 i^2 s_1^2 - \alpha_1 a^2 i^2 s_2^2 + s_2^2 \tilde{r}(i) \right) \quad (65)$$

$$= \frac{(1-p_1)}{(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} p_1^{i-1} i^2 s^2 + s_2^2 \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} (\tilde{r}(i) - \alpha_1 a^2 i^2) p_1^{i-1}. \quad (66)$$

Then,

$$\sum_{i,j} \pi_{i,j} \tilde{z}_{i,j} = v \frac{(1-p_1)}{(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} i p_1^{i-1} + \sum_{i=1}^{\alpha_1} \sum_{j=1}^{\alpha_1} \pi_{i,j} \tilde{z}_{i,j}^2 \quad (67)$$

$$= \frac{(1-p_1)}{(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} [iv + i^2 s^2] p_1^{i-1} + s_2^2 \frac{(1-p_1)}{\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} (\tilde{r}(i) - \alpha_1 a^2 i^2) p_1^{i-1}. \quad (68)$$

Substituting (58) and (68) in (9) yields $\mathbb{E}[\Delta_1]$, $\mathbb{E}[\Delta_2]$, can be obtained by replacing a , s , v , s_1 , s_2 , p_1 , and α_1 in (12) with $\frac{1}{a}$, $\frac{s}{a}$, $\frac{v}{a}$, s_2 , s_1 , p_2 , and α_2 , respectively.

APPENDIX D PROOF OF THEOREM 3

Note that to minimize (12), for a fixed α_1 and α_2 , it is sufficient to look at the following optimization problem.

$$\begin{aligned} \min_{r_i \in \mathbb{Z}_0^+} & \sum_{i=1}^{\alpha_1} \tilde{r}(i) p_1^{i-1} \\ \text{s.t.} & \sum_{i=1}^{\alpha_1} r_i = \alpha_2. \end{aligned} \quad (69)$$

Here, $\tilde{r}(i)$ can be viewed simply as the sum of squares of the sum of i consecutive terms in the placement vector. For example, $\tilde{r}(1)$, $\tilde{r}(2)$ and $\tilde{r}(3)$ are,

$$\tilde{r}(1) = r_1^2 + r_2^2 + \dots + r_{\alpha_1}^2, \quad (70)$$

$$\tilde{r}(2) = (r_1 + r_2)^2 + (r_2 + r_3)^2 + \dots + (r_{\alpha_1} + r_1)^2, \quad (71)$$

$$\begin{aligned} \tilde{r}(3) &= (r_1 + r_2 + r_3)^2 + (r_2 + r_3 + r_4)^2 + \dots \\ &\quad + (r_{\alpha_1} + r_1 + r_2)^2. \end{aligned} \quad (72)$$

Now, to solve the above optimization problem, we look at the following sub-problem.

$$\begin{aligned} \min_{r_i \in \mathbb{Z}_0^+} & \tilde{r}(i) \\ \text{s.t.} & \sum_{i=1}^{\alpha_1} r_i = \alpha_2. \end{aligned} \quad (73)$$

Denote this sub-problem by $J(i)$ and its optimal value by $J^*(i)$. Let $C_r(k)$ be a counting function that gives the number of elements in the placement vector $\mathbf{r}$ with value k . First, we will find the structure of the optimal placement vector that minimizes $J(1)$.

Lemma 1: Let r_i^* denote the elements of the optimal placement vector that minimizes $J(1)$. Then, $r_i^* \in \{[a], \lfloor a \rfloor\}$ where $a = \frac{\alpha_2}{\alpha_1}$. Moreover, if $a \notin \mathbb{Z}$, then $C_{r^*}(\lceil a \rceil) = \alpha_1(a - \lfloor a \rfloor)$ and $C_{r^*}(\lfloor a \rfloor) = \alpha_1(\lceil a \rceil - a)$.

Proof: Suppose a is an integer. Now, if we solve the optimization problem by relaxing the integer constraints, it turns out that $r_i = a$ is its unique minimizer. Since a is already an integer, $r_i = a$ will be the optimal for the integer program as well.

Now, suppose $a \notin \mathbb{Z}$. Therefore, $\alpha_1 \lfloor a \rfloor < \alpha_2 < \alpha_1 \lceil a \rceil$. Assume for some i , $r_i^* > \lceil a \rceil$. Then, $\exists r_j^*$ such that $r_j \leq \lfloor a \rfloor$. Now, let us look at the following placement vector λ where $\lambda_k = r_k^*$ for $k \neq i \neq j$ with $\lambda_i = r_i^* - 1 \geq 0$ and $\lambda_j = r_j^* + 1$. Note that $\sum_k \lambda_k = \alpha_2$ and $\sum_k \lambda_k^2 = 2(1 + r_j - r_i) + \sum_k r_k^* < J^*(1)$. This is a contradiction, and hence $r_i^* \leq \lceil a \rceil, \forall i$. Similarly, we can prove $r_i^* \geq \lfloor a \rfloor, \forall i$.

Now, to find $C_r(\lceil a \rceil)$ and $C_r(\lfloor a \rfloor)$, we simply need to solve the following equations,

$$\lceil a \rceil C_{r^*}(\lceil a \rceil) + \lfloor a \rfloor C_{r^*}(\lfloor a \rfloor) = \alpha_2, \quad (74)$$

$$C_{r^*}(\lceil a \rceil) + C_{r^*}(\lfloor a \rfloor) = \alpha_1. \quad (75)$$

This will yield that $C_{r^*}(\lceil a \rceil) = \alpha_1(a - \lfloor a \rfloor)$ and $C_{r^*}(\lfloor a \rfloor) = \alpha_1(\lceil a \rceil - a)$. ■

Next, we try to find the optimal placement vector for $J(i)$. If a is an integer, then relaxing integer constraints yields that $r_i = a, \forall i$ is a minimizer for $J(i)$ and hence is optimal for the original integer program as well. Thus, for the rest of this proof, we will assume that $a \notin \mathbb{Z}$. Now, to minimize $J(i)$, we look at the following relaxed version of the problem,

$$\begin{aligned} \min_{z_k \in \mathbb{Z}_0^+} & \sum_{k=1}^{\alpha_1} (z_{(k-1)i+1} + z_{(k-1)i+2} + \dots + z_{ki})^2 \\ \text{s.t.} & \sum_{k=1}^{i\alpha_1} z_k = i\alpha_2. \end{aligned} \quad (76)$$

Let this optimization problem be denoted by $\tilde{J}(i)$ and its optimal value by $\tilde{J}^*(i)$. Note that if we add the constraints $z_{(k-1)i+j} = z_{ki+j-1}$ for $j > 2, 1 \leq k < \alpha_1$ and $z_{(\alpha_1-1)i+j} = z_{j-1}$ for $j > 2$ to $\tilde{J}(i)$, we get $J(i)$. Hence, $J^*(i) \geq \tilde{J}^*(i)$.

Lemma 2: Let z_k^* denote the optimal integers that minimize $\tilde{J}(i)$. Then, there exists an optimal solution where $z_k^* \in \{\lceil a \rceil, \lfloor a \rfloor\}$. Moreover, let $b = \alpha_1(a - \lfloor a \rfloor)$ and $c = \alpha_1(\lceil a \rceil - a)$. Then, $C_{z^*}(\lceil a \rceil) = ib$ and $C_{z^*}(\lfloor a \rfloor) = ic$.

Proof: Let $y_k = z_{(k-1)i+1} + z_{(k-1)i+2} + \dots + z_{ki}$. Now, consider the following optimization problem,

$$\begin{aligned} \min_{y_k \in \mathbb{Z}_0^+} \quad & \sum_{k=1}^{\alpha_1} y_k^2 \\ \text{s.t.} \quad & \sum_{k=1}^{\alpha_1} y_k = i\alpha_2. \end{aligned} \quad (77)$$

Denote the above optimization problem by $\hat{J}(i)$ and its optimal solution by y_k^* . From Lemma 1, we have that $y_k^* \in \left\{ \lceil \frac{i\alpha_2}{\alpha_1} \rceil, \lfloor \frac{i\alpha_2}{\alpha_1} \rfloor \right\}$. Now, from (74), we have that $\alpha_2 = c\lfloor a \rfloor + b\lceil a \rceil = \alpha_1\lfloor a \rfloor + b$. Therefore, $i\alpha_2 = i\alpha_1\lfloor a \rfloor + ib$. This yields that,

$$\lceil \frac{i\alpha_2}{\alpha_1} \rceil = i\lfloor a \rfloor + \lceil \frac{ib}{\alpha_1} \rceil = \left(i - \lceil \frac{ib}{\alpha_1} \rceil \right) \lfloor a \rfloor + \lceil \frac{ib}{\alpha_1} \rceil \lceil a \rceil, \quad (78)$$

$$\lfloor \frac{i\alpha_2}{\alpha_1} \rfloor = i\lfloor a \rfloor + \lfloor \frac{ib}{\alpha_1} \rfloor = \left(i - \lfloor \frac{ib}{\alpha_1} \rfloor \right) \lfloor a \rfloor + \lfloor \frac{ib}{\alpha_1} \rfloor \lceil a \rceil. \quad (79)$$

This means that each y_k^* can be constructed using $z_{(k-1)i+j} \in \{\lceil a \rceil, \lfloor a \rfloor\}$. Further, since $\hat{J}(i)$ is a relaxed version of $\tilde{J}(i)$, an optimal solution for $\tilde{J}(i)$ can be constructed solely with $z_k^* \in \{\lceil a \rceil, \lfloor a \rfloor\}$.

Now, to prove that $C_{z^*}(\lceil a \rceil) = ib$ and $C_{z^*}(\lfloor a \rfloor) = ic$, if $\frac{i\alpha_2}{\alpha_1} \notin \mathbb{Z}$ again, from Lemma 1, we have,

$$C_{y^*} \left(\lceil \frac{i\alpha_2}{\alpha_1} \rceil \right) = i\alpha_2 - \alpha_1 \lfloor \frac{i\alpha_2}{\alpha_1} \rfloor = ib - \alpha_1 \lfloor \frac{ib}{\alpha_1} \rfloor, \quad (80)$$

$$C_{y^*} \left(\lfloor \frac{i\alpha_2}{\alpha_1} \rfloor \right) = \alpha_1 \lceil \frac{i\alpha_2}{\alpha_1} \rceil - i\alpha_2 = \alpha_1 \lceil \frac{ib}{\alpha_1} \rceil - ib. \quad (81)$$

Therefore, we have,

$$C_{z^*}(\lceil a \rceil) = C_{y^*} \left(\lceil \frac{i\alpha_2}{\alpha_1} \rceil \right) \lceil \frac{ib}{\alpha_1} \rceil + C_{y^*} \left(\lfloor \frac{i\alpha_2}{\alpha_1} \rfloor \right) \lfloor \frac{ib}{\alpha_1} \rfloor \lceil a \rceil \quad (82)$$

$$= \left(ib - \alpha_1 \lfloor \frac{ib}{\alpha_1} \rfloor \right) \lceil \frac{ib}{\alpha_1} \rceil + \left(\alpha_1 \lceil \frac{ib}{\alpha_1} \rceil - ib \right) \lfloor \frac{ib}{\alpha_1} \rfloor \lceil a \rceil \quad (83)$$

$$= ib. \quad (84)$$

Hence, $C_{z^*}(\lfloor a \rfloor) = i\alpha_1 - C_{z^*}(\lceil a \rceil) = ic$. If $\frac{i\alpha_2}{\alpha_1} \in \mathbb{Z}$, then $C_{z^*}(\lceil a \rceil) = \alpha_1 \lceil \frac{ib}{\alpha_1} \rceil = ib$. ■

From Lemma 2 we have that there exists an optimal placement z^* that minimizes $\tilde{J}(i)$, whose elements are either $\lfloor a \rfloor$ or $\lceil a \rceil$. Let r^* be the minimizer of $J(1)$ where $\lfloor a \rfloor$ and $\lceil a \rceil$ elements are almost uniformly distributed within the vector. Now, to construct z^* , we consider the following two sub cases.

- **i and α_1 are coprime:** In here, z^* can be constructed by distributing the ib $\lceil a \rceil$ s and ic $\lfloor a \rfloor$ s as uniformly as possible within the z^* vector so that the sum of the i terms in each of the α_1 windows is as uniform as possible. This is equivalent to setting $z^* = \{r^*, r^*, \dots, r^*\}$ by repeating r^* i times since r^* is composed of b $\lceil a \rceil$ terms and c $\lfloor a \rfloor$ terms. Now, let us look at how the objective function would behave in this z^* . Note that if i and α_1 are coprime, then $ik \bmod \alpha_1$ would cycle through the values from 0 to $\alpha_1 - 1$ as k goes from 1 to α_1 . This means that by evaluating $\tilde{J}(i)$ for z^* , we will essentially be evaluating

$J(i)$ for r^* . To see this, let us consider a simple example where $\alpha_1 = 5$ and $i = 3$. Then, z^* is as follows,

$$\begin{aligned} z^* &= \{r_1^*, r_2^*, r_3^*, r_4^*, r_5^*, \\ &\quad r_1^*, r_2^*, r_3^*, r_4^*, r_5^*, \\ &\quad r_1^*, r_2^*, r_3^*, r_4^*, r_5^*\} \end{aligned} \quad (85)$$

$$\begin{aligned} &= \{r_1^*, r_2^*, r_3^*, \\ &\quad r_4^*, r_5^*, r_1^*, \\ &\quad r_2^*, r_3^*, r_4^*, \\ &\quad r_5^*, r_1^*, r_2^*, \\ &\quad r_3^*, r_4^*, r_5^*\}. \end{aligned} \quad (86)$$

As seen by (86), then the i partitions of z^* corresponds to the cyclic sum of 3 consecutive elements in r^* . Now, since $J^*(i) \geq \tilde{J}^*(i)$ and z^* achieves $\tilde{J}^*(i)$, it turns out r^* achieves $J^*(i)$.

- **$\gcd(\alpha_1, i) = d > 1$:** Let $\text{cyc}(r^*, j)$ be vector that is formed by cyclically shifting the vector r^* to the right by j elements. Now, we can construct an equally optimal z^* as follows,

$$\begin{aligned} z^* &= \left\{ \underbrace{r^*, \dots, r^*}_{\frac{i}{d} \text{ times}}, \underbrace{\text{cyc}(r^*, 1), \dots, \text{cyc}(r^*, 1)}_{\frac{i}{d} \text{ times}}, \right. \\ &\quad \left. \dots, \underbrace{\text{cyc}(r^*, d-1), \dots, \text{cyc}(r^*, d-1)}_{\frac{i}{d} \text{ times}} \right\}. \end{aligned} \quad (87)$$

Now, similar to the coprime case, if we evaluate $\tilde{J}(i)$ for z^* we will be essentially evaluating $J(i)$ for r^* . Hence, r^* is optimal for $J(i)$ as well. To see this, as before let us look at a simple example where $\alpha_1 = 6$ and $i = 2$. Here $d = 2$. Then, z^* will be as follows,

$$\begin{aligned} z^* &= \{r_1^*, r_2^*, r_3^*, r_4^*, r_5^*, r_6^*, \\ &\quad r_2^*, r_3^*, r_4^*, r_5^*, r_6^*, r_1^*\} \end{aligned} \quad (88)$$

$$\begin{aligned} &= \{r_1^*, r_2^*, \\ &\quad r_3^*, r_4^*, \\ &\quad r_5^*, r_6^*, \\ &\quad r_2^*, r_3^*, \\ &\quad r_4^*, r_5^*, \\ &\quad r_6^*, r_1^*\}. \end{aligned} \quad (89)$$

As seen by (89), then the i partitions of z^* corresponds to the cyclic sum of 2 consecutive elements in r^* .

Now, since the almost uniform r^* is optimal for all $J(i)$, it is optimal for our original problem.

APPENDIX E PROOF OF THEOREM 4

In this section, we will prove that for all $\epsilon > 0$, the minimum weighted AoI obtained by the solution of Algorithm 2 can be made to be within ϵ of the actual optimum by choosing α in Algorithm 2 sufficiently large. Let us analyze the age

expression of source-1. Define the functions $f(a)$, $g(a)$, and $h(a)$ for $a = \frac{\alpha_2}{\alpha_1}$ as follows,

$$f(a) = \frac{(1+p_1)}{2(1-p_1)}s + \frac{v}{2s} + s_1, \quad (90)$$

$$g(a) = a^2 s_2^2 \frac{(1-p_1)^2}{2s(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} i^2 p_1^{i-1}, \quad (91)$$

$$h(a) = s_2^2 \frac{(1-p_1)^2}{2s\alpha_1(1-p_1^{\alpha_1})} \sum_{i=1}^{\alpha_1} \tilde{r}(i) p_1^{i-1}. \quad (92)$$

Then, $\mathbb{E}[\Delta_1]$ given in (12) is equivalent to,

$$\mathbb{E}[\Delta_1] = f(a) - g(a) + h(a). \quad (93)$$

Let the optimal a be denoted by $a^* = \frac{\alpha_2^*}{\alpha_1^*} \notin \mathbb{Z}$ (if $a^* \in \mathbb{Z}$, the result is trivial) where α_1^* and α_2^* are coprime. Without loss of generality assume $a^* > 1$. Now, partition the interval from $[1, a_{\max}]$ into equal partitions of length $\frac{1}{\alpha}$. Trivially, a^* must be present in one of these partitions. If a partition end point coincides with a^* , then we will have the optimum solution. Let us analyze the case where a^* falls within a partition. Then, $\exists \alpha_2$ such that $\alpha_2 > \alpha_1 = \alpha$ and $\lfloor a^* \rfloor \leq \frac{\alpha_2-1}{\alpha_1} < \frac{\alpha_2^*}{\alpha_1^*} < \frac{\alpha_2}{\alpha_1} \leq \lceil a^* \rceil$. Therefore, the placement vector of pattern (α_1, α_2) will be composed of $\lfloor a^* \rfloor$ and $\lceil a^* \rceil$ terms only. We will now try to bound the differences of functions f , g and h at values a^* and $a = \frac{\alpha_2}{\alpha_1}$. As described in Section V-A, the schedule (α_1^*, α_2^*) is equivalent to the schedule $(k\alpha_1^*, k\alpha_2^*)$. Therefore, we will be analyzing these functions by treating $a^* = \frac{\alpha_1\alpha_2^*}{\alpha_1^*\alpha_1}$ and $a^* = \frac{\alpha_1^*\alpha_2}{\alpha_1^*\alpha_1}$. In particular, we will bound the AoI difference of patterns $(\alpha_1\alpha_2^*, \alpha_1\alpha_1^*)$ and $(\alpha_1^*\alpha_2, \alpha_1^*\alpha_1)$. For brevity, let us denote these two patterns as P^* and P .

Let $h_1(a) = \frac{s_2^2}{2s}$. Since $f(a)$ is a continuous function of a , $\exists \delta_1 > 0$ such that $|a^* - a| < \delta_1 \implies |f(a) - f(a^*)| < \epsilon$. Since $h_1(a)$ is a continuous function of a , $\exists \delta_2$ such that $|a^* - a| < \delta_2 \implies |h_1(a) - h_1(a^*)| < \epsilon(1-p_1)$. Moreover, since $h_1(a)$ is continuous, $h_1(a)$ is bounded on the interval $[a_{\min}, a_{\max}]$ defined by the RR policy. Let $h_1(a) < M_1$ for $a \in [a_{\min}, a_{\max}]$. Also, as $\alpha_1 \rightarrow \infty$, $\sum_{i=1}^{\alpha_1} i^2 p_1^{i-1} / (1-p_1^{\alpha_1})$ converges to $\sum_{i=1}^{\infty} i^2 p_1^{i-1} = M < \infty$. Therefore $\exists \delta_3 > 0$ such that $|a^* - a| < \delta_3 \implies |g(a) - g(a^*)| < \epsilon$. Moreover, note that $\sum_{i=\alpha_0}^{\alpha_1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} < \sum_{i=\alpha_0}^{\alpha_1} i^2 [a]^2 p_1^{i-1} < \sum_{i=\alpha_0}^{\infty} i^2 [a]^2 p_1^{i-1} < [a]^2 M$. Therefore, $\exists \alpha_0$ such that $\alpha_1 > \alpha_0 \implies \sum_{i=\alpha_0}^{\alpha_1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} < \frac{\epsilon}{2}$. At the same time, this α_0 can be chosen such that $\alpha_1 > \alpha_0 \implies \frac{1}{1-p_1^{\alpha_1}} < 1 + \frac{\epsilon}{2[a]^2 M}$. This means that the following inequalities hold for $\alpha_1 > \alpha_0$,

$$\begin{aligned} \sum_{i=1}^{\alpha_0-1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} - \epsilon &< \frac{1}{1-p_1^{\alpha_1}} \sum_{i=1}^{\alpha_1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} \\ &< \left(1 + \frac{\epsilon}{2[a]^2 M}\right) \left(\sum_{i=1}^{\alpha_0-1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} + \frac{\epsilon}{2}\right), \end{aligned} \quad (94)$$

$$\begin{aligned} \sum_{i=1}^{\alpha_0-1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} - \epsilon &< \frac{1}{1-p_1^{\alpha_1}} \sum_{i=1}^{\alpha_1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} \\ &< \sum_{i=1}^{\alpha_0-1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1} + \epsilon. \end{aligned} \quad (95)$$

Let $h_2(a) = \frac{1}{1-p_1^{\alpha_1}} \sum_{i=1}^{\alpha_1} \frac{\tilde{r}(i)}{\alpha_1} p_1^{i-1}$. Let us represent the $\tilde{r}(i)$ with respect to the optimal placement vector of P^* as $\tilde{r}_1(i)$ and that

of P as $\tilde{r}_2(i)$. From Lemma 1, it follows that the placement vector for P would have $\alpha_2\alpha_1^* - \alpha_1\alpha_2^*$ extra $\lceil a^* \rceil$ than that of P^* . Therefore, $\tilde{r}_2(i) \geq \tilde{r}_1(i)$. Then, from (95), we have that,

$$|h_2(a) - h_2(a^*)| \leq 2\epsilon + \sum_{i=1}^{\alpha_0-1} \frac{\tilde{r}_2(i) - \tilde{r}_1(i)}{\alpha_1\alpha_1^*} p_1^{i-1}. \quad (96)$$

Note that since a^* belongs to a partition, we have that $a - a^* = \frac{\alpha_2\alpha_1^* - \alpha_1\alpha_2^*}{\alpha_1\alpha_1^*} < \frac{1}{\alpha_1}$. This condition tells us that if we divide the placement vector of P^* into segments of α_1 terms and increase one of the terms in these segments by one, we can construct a suboptimal placement vector for P . Further, if $\alpha_1 > 2\alpha_0$, these modifications can be done such that none of the modified terms fall within a window of α_0 elements. Let the $\tilde{r}(i)$ for this suboptimal placement vector be denoted by $\tilde{r}_3(i)$. Since it is a suboptimal vector, we have $\tilde{r}_2(i) \leq \tilde{r}_3(i)$. Moreover, $\tilde{r}_3(i) - \tilde{r}_1(i) \leq (\alpha_2\alpha_1^* - \alpha_1\alpha_2^*)(2i)(\lceil a^* \rceil + 1)^2$ since we have $\alpha_2\alpha_1^* - \alpha_1\alpha_2^*$ modifications and each modification would affect $2i$ summation windows where each summation window is bounded by $i(\lceil a^* \rceil + 1)$. Therefore we have,

$$\begin{aligned} |h_2(a) - h_2(a^*)| &\leq 2\epsilon + \frac{2(\alpha_2\alpha_1^* - \alpha_1\alpha_2^*)(\lceil a^* \rceil + 1)^2}{\alpha_1\alpha_1^*} \sum_{i=1}^{\alpha_0-1} i^3 p_1^{i-1} \end{aligned} \quad (97)$$

$$\leq 2\epsilon + \frac{2(\lceil a_{\max} \rceil + 1)^2}{\alpha_1} \sum_{i=1}^{\alpha_0-1} i^3 p_1^{i-1}. \quad (98)$$

Then, from (98), $\exists k_0$ such that $\alpha_1 > k_0 \implies |h_2(a) - h_2(a^*)| < 3\epsilon$. Now, select $\alpha_1 = \alpha > \max\{k_0, 2\alpha_0\}$. Further, if α is such that $\frac{1}{\alpha} < \min\{\delta_1, \delta_2, \delta_3\}$ we have that $|a - a^*| < \min\{\delta_1, \delta_2, \delta_3\}$. Then, the following inequalities holds true,

$$\begin{aligned} |h(a) - h(a^*)| &< h_1(a)|h_2(a) - h_2(a^*)| + h_2(a^*)|h_1(a) - h_1(a^*)| \end{aligned} \quad (99)$$

$$\begin{aligned} &\leq M_1|h_2(a) - h_2(a^*)| \\ &\quad + \frac{\lceil a_{\max} \rceil^2}{1-p_1} \sum_{i=1}^{\alpha_1\alpha_1^*} i^2 p_1^{i-1} |h_1(a) - h_1(a^*)| \end{aligned} \quad (100)$$

$$\leq 3M_1\epsilon + \frac{M\lceil a_{\max} \rceil^2}{1-p_1} |h_1(a) - h_1(a^*)| \quad (101)$$

$$< (3M_1 + M\lceil a_{\max} \rceil^2)\epsilon. \quad (102)$$

Therefore, we have that $|(f(a) - g(a) + h(a)) - (f(a^*) - g(a^*) + h(a^*))| < |f(a) - f(a^*)| + |g(a) - g(a^*)| + |h(a) - h(a^*)| < (3M_1 + M\lceil a_{\max} \rceil^2 + 2)\epsilon$. Similarly, we can bound the difference of $\mathbb{E}[\Delta_2]$ at a and a^* as well. Therefore, by selecting α large enough we can get a cyclic schedule whose weighted AoI is closer to the weighted AoI of the optimal cyclic scheduler. This is equivalent to setting α in Algorithm 2 to a large value and either increasing the numerator (for $a^* > 1$) by fixing the denominator as α or by increasing the denominator (for $a^* < 1$) by fixing the numerator as α . Note that, if a function is continuous in a compact interval, then it is uniformly continuous on that interval. Hence, $\delta_1, \delta_2, \delta_3$ can be chosen independent of a^* . Moreover, α_0 and k_0 are independent of a^* . Therefore, we can obtain α_ϵ independently of a^* .

APPENDIX F

MINIMIZATION OF EXPRESSION (22)

We are interested in finding the source utilization parameters τ_n^* that will minimize the expression for weighted AoI given in (22) for given $\tilde{c}_n$ values. This leads us to the following optimization problem,

$$\begin{aligned} \min_{\tau_n \geq 0} \quad & \sum_{n=1}^N \left(a_n \tau_n + b_n \frac{1}{\tau_n} \right) \\ \text{s.t.} \quad & \sum_{n=1}^N \tau_n = 1, \end{aligned} \quad (103)$$

where the coefficients a_n and b_n are non-negative. The optimization problem (103) is convex and the KKT conditions result in the following condition,

$$a_n - \frac{b_n}{\tau_n^2} = a_m - \frac{b_m}{\tau_m^2}, \quad 1 \leq n, m \leq N, \quad (104)$$

which together with the normalization constraint in (103) yields the following non-linear fixed point equation in the single unknown $x \in \mathbb{R}$,

$$f_p(x) = \sum_{n=1}^N \sqrt{\frac{b_n}{a_n - x}} - 1 = 0. \quad (105)$$

We note that the function $f_p(x) \in (-1, \infty)$ is a monotonically increasing function of $x \in (-\infty, \min_{1 \leq n \leq N} a_n)$, and therefore, has a unique solution denoted by x^* (can be found by bisection search) from which one can find the optimum coefficients τ_n^* , $1 \leq n \leq N$ as follows,

$$\tau_n^* = \sqrt{\frac{b_n}{a_n - x^*}}. \quad (106)$$

APPENDIX G

CHARACTERISTICS OF ALGORITHM 5

In this section, we present a few properties of the packet spreading algorithm presented in Algorithm 5 that ensures the generation of a cyclic schedule which adheres to the number of occurrences $\{K_n\}_{n=1}^N$ of each source given by Algorithm 4. In particular, we would like to show that the output pattern would contain exactly K_n occurrences of source- n and all the deficit counters would be zero at the end of K iterations.

Let B_n^k denote the value of the deficit counter for the n th source in the k th iteration of Algorithm 5. Let Q_n^k denote the quantum allocated for source- n in the k th iteration, where $Q_n^k = \frac{(1-B_n^k)K}{K_n}$. Since the initial deficit counters are all set to zero, $Q_n^0 = \frac{1}{K_n}$. In each iteration, the source with the smallest quantum is selected as the source to be inserted into the pattern P . Then, the deficit counter of the selected source is set to zero again and the deficit counters of all the other sources are updated by the equation $B_n^{k+1} = B_n^k + Q_{\tilde{Q}_j^k}$, where $Q = \min_n Q_n^k$. Based on the above steps, Lemma 3 gives the update equations for the quantum allocated for each source.

Lemma 3: The quantum Q_n^k allocated for source- n is updated according to the equation,

$$Q_n^{k+1} = \begin{cases} Q_n^k - \min_n Q_n^k, & n \neq \arg \min_j Q_j^k, \\ \frac{K}{K_n}, & n = \arg \min_j Q_j^k. \end{cases} \quad (107)$$

Proof: If $n \neq \arg \min_j Q_j^k$, then, the quantum in the next iteration is given by,

$$Q_n^{k+1} = \frac{(1 - B_n^{k+1})K}{K_n} \quad (108)$$

$$= \frac{(1 - B_n^k - Q_{\tilde{Q}_j^k})K}{K_n} \quad (109)$$

$$= \frac{(1 - B_n^k)K}{K_n} - Q \quad (110)$$

$$= Q_n^k - \min_n Q_n^k. \quad (111)$$

If $n = \arg \min_j Q_j^k$, then, $B_n^{k+1} = 0$. Thus, $Q_n^{k+1} = \frac{K}{K_n}$. ■

Since all Q_n^k s have a scaling factor of K , we can safely normalize all the Q_n^k by K . Let $\tilde{Q}_n^k$ represent the normalized Q_n^k values. Then, the source to be inserted in each iteration is given by $\arg \min_j \tilde{Q}_j^k$, and its update equations are given by,

$$\tilde{Q}_n^{k+1} = \begin{cases} \tilde{Q}_n^k - \min_n \tilde{Q}_n^k, & n \neq \arg \min_j \tilde{Q}_j^k, \\ \frac{1}{K_n}, & n = \arg \min_j \tilde{Q}_j^k. \end{cases} \quad (112)$$

We will first prove the desired properties for $N = 2$, and later show how it can be extended for $N > 2$. Without loss of generality, assume $K_1 \geq K_2 \geq \dots \geq K_N$. Let x_n^k represent the number of source- n instances that had been inserted into the pattern after the k th iteration.

Property 1: For $N = 2$, $\tilde{Q}_{N-i+1}^{k+1} = \frac{x_{N-i+1}^k + 1}{K_{N-i+1}} - \frac{x_i^k}{K_i}$, where i is the source that was selected in the k th iteration.

Proof: Initially, the pattern P is empty with $\tilde{Q}_1^0 = \frac{1}{K_1}$ and $\tilde{Q}_2^0 = \frac{1}{K_2}$. Since $K_1 \geq K_2$, $\tilde{Q}_1^0 \leq \tilde{Q}_2^0$. Therefore, we would insert a source-1 scheduling instance to the pattern P . Updating the normalized quantum would yield that $\tilde{Q}_1^1 = \frac{1}{K_1}$ and $\tilde{Q}_2^1 = \frac{1}{K_2} - \frac{1}{K_1}$. Note that Property 1 is true for $k = 0$. We will now prove that the result will hold for $k = m + 1$ assuming that it holds for $k = m$. For brevity, let $j = N - i + 1$, which is the source that was not selected in the m th iteration. Based on our assumption $\tilde{Q}_j^{m+1} = \frac{x_j^m + 1}{K_j} - \frac{x_i^m}{K_i}$, where i is the source that was selected in the m th iteration. Since i was selected in the m th iteration, $\tilde{Q}_i^{m+1} = \frac{1}{K_i}$. There are two possibilities to consider. If source- i was selected again in the $(m + 1)$ th iteration, then $x_j^{m+1} = x_j^m$ and $x_i^{m+1} = x_i^m + 1$. Therefore, $\tilde{Q}_j^{m+2} = \frac{x_j^m + 1}{K_j} - \frac{x_i^m + 1}{K_i} = \frac{x_j^m + 1}{K_j} - \frac{x_i^{m+1}}{K_i}$. This satisfies Property 1. If source- j was selected in the $(m + 1)$ th iteration, then $x_j^{m+1} = x_j^m + 1$ and $x_i^{m+1} = x_i^m$. Therefore, $\tilde{Q}_i^{m+2} = \frac{x_i^m + 1}{K_i} - \frac{x_j^m + 1}{K_j} = \frac{x_i^{m+1}}{K_i} - \frac{x_j^{m+1}}{K_j}$. This too satisfies Property 1. Therefore by induction, Property 1 holds for all $k \geq 0$. ■

Property 2: For $N = 2$, $x_n^{K-1} = K_n$.

Proof: Since $\tilde{Q}_1^0 \leq \tilde{Q}_2^0$, we would insert source-1 scheduling instances into the pattern until $\tilde{Q}_2^k < \frac{1}{K_1}$. Suppose we insert a source-2 scheduling instance for the first time in the m_1 th iteration. Then, $\tilde{Q}_2^{m_1} = \frac{1}{K_2} - \frac{m_1}{K_1} < \tilde{Q}_1^{m_1} = \frac{1}{K_1}$. Updating the normalized quantum yields, $\tilde{Q}_2^{m_1+1} = \frac{1}{K_2}$ and $\tilde{Q}_1^{m_1+1} = \frac{m_1+1}{K_1} - \frac{1}{K_2}$. Here, m_1 would correspond to the smallest positive integer such that $\frac{m_1+1}{K_1} - \frac{1}{K_2} > 0$. This implies that $\frac{m_1}{K_1} - \frac{1}{K_2} < 0$ and

$\tilde{Q}_1^{m_1+1} < \frac{1}{K_1} \leq \frac{1}{K_2} = \tilde{Q}_2^{m_1+1}$. Therefore, a source-1 instance would be inserted to the pattern in the $(m_1 + 1)$ th iteration. Updating the normalized quantum again would yield that $\tilde{Q}_1^{m_1+2} = \frac{1}{K_1}$ and $\tilde{Q}_2^{m_1+2} = \frac{2}{K_2} - \frac{m_1+1}{K_1}$. Then, again we will be inserting source-1 scheduling instances into the pattern until $\frac{2}{K_2} - \frac{m_2}{K_1} < \frac{1}{K_1}$. Here, $m_2 + 1$ is the iteration at which we insert the second instance for source-2 into the pattern where m_2 corresponds to the smallest positive integer such that $\frac{m_2+1}{K_1} - \frac{2}{K_2} > 0$. Moreover, m_2 also corresponds to the number of source-1 instances inserted into the pattern. Extending this argument K_2 times, we get that K_2 th scheduling instance of source-2 will be inserted into the pattern when $\frac{K_2}{K_2} - \frac{m_{K_2}}{K_1} < \frac{1}{K_1}$, where m_{K_2} is the smallest positive integer that satisfies the inequality and $m_{K_2} + K_2 - 1$ is the iteration at which the inequality is satisfied. This gives us that $m_{K_2} = K_1$, and therefore, the iteration where we insert the K_2 th scheduling instance of source-2 is $K_1 + K_2 - 1 = K - 1$, proving the desired result. ■

Property 3: For $N > 1$, $\tilde{Q}_j^{k+1} = \frac{x_j^{k+1}}{K_j} - \frac{x_i^k}{K_i}$, where i is the source that was selected in the k th iteration and $j \neq i$.

Property 3 is proved using an induction argument similar to Property 1.

Property 4: For $N > 1$, $x_n^{K-1} = K_n$ and $\tilde{Q}_n^K = \frac{1}{K_n}$. Therefore, $B_n^K = 0$.

Proof: Consider source- j and source- N together. Note that, whenever we select another source other than source- j or source- N to be inserted into the pattern, we will be subtracting the same constant from both $\tilde{Q}_j^k$ and $\tilde{Q}_N^k$. Therefore, the insertion of source- j relative to source- N insertions would follow the same pattern when the other sources are considered to be absent. Therefore, from Property 2, by the time we insert the K_N th scheduling instance of source- N into the pattern we would have inserted K_j instances of source- j into the pattern. This is true for all $1 \leq j < N$. Therefore, the K_N th instance of source- N would be inserted into the pattern in the iteration $K - 1$, and thus $x_j^{K-1} = K_j$. Since we selected source- N in the $(K - 1)$ th iteration, we have $\tilde{Q}_N^K = \frac{1}{K_N}$. From Property 3, we have that $\tilde{Q}_j^K = \frac{x_j^{K-1} + 1}{K_j} - \frac{x_N^{K-1}}{K_N} = \frac{K_j + 1}{K_j} - \frac{K_N}{K_N} = \frac{1}{K_j}$. This implies that $B_n^K = 0$. ■

ACKNOWLEDGMENT

This work is done when Nail Akar is on sabbatical leave as a visiting professor at University of Maryland, MD, USA, which is supported in part by the Scientific and Technological Research Council of Türkiye (Tübitak) 2219-International Postdoctoral Research Fellowship Program.

REFERENCES

- [1] Y. Gu, H. Chen, Y. Zhou, Y. Li, and B. Vucetic, "Timely status update in Internet of Things monitoring systems: An age-energy tradeoff," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5324–5335, Jun. 2019.
- [2] R. A. Gupta and M.-Y. Chow, "Networked control system: Overview and research trends," *IEEE Trans. Ind. Electron.*, vol. 57, no. 7, pp. 2527–2535, Jul. 2009.
- [3] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," in *Proc. 10th ACM/IEEE Int. Conf. Cyber-Physical Syst.*, Apr. 2019, pp. 109–117.
- [4] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1807–1827, Mar. 2019.
- [5] S. K. Kaul, R. D. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. IEEE Infocom*, Mar. 2012, pp. 2731–2735.
- [6] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1183–1210, May 2021.
- [7] A. Kosta, N. Pappas, and V. Angelakis, *Age of Information: A New Concept, Metric and Tool*. Boston, MA, USA: Now, 2017.
- [8] Y. Sun, I. Kadota, R. Talak, and E. Modiano, "Age of information: A new metric for information freshness," *Synth. Lectures Commun. Netw.*, vol. 12, no. 2, pp. 1–224, Dec. 2019.
- [9] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 11891–11915, Aug. 2021.
- [10] M. Jouhari, N. Saeed, M.-S. Alouini, and E. M. Amhoud, "A survey on scalable LoRaWAN for massive IoT: Recent advances, potentials, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 3, pp. 1841–1876, 3rd Quart., 2023.
- [11] I. Kadota, M. S. Rahman, and E. Modiano, "WiFresh: Age-of-Information from theory to implementation," in *Proc. Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2021, pp. 1–11.
- [12] J. Liu, R. Zhang, A. Gong, and H. Chen, "Optimizing age of information in wireless uplink networks with partial observations," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 4105–4118, Jul. 2023.
- [13] B. Yu, Y. Cai, D. Wu, C. Dong, R. Zhang, and W. Wu, "Optimizing age of information for uplink cellular Internet of Things with random access," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 20300–20313, Jun. 2024.
- [14] R. D. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 3008–3012.
- [15] H. Chen, Y. Gu, and S.-C. Liew, "Age-of-information dependent random access for massive IoT networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Jul. 2020, pp. 930–935.
- [16] I. Kadota and E. Modiano, "Age of information in random access networks with stochastic arrivals," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, May 2021, pp. 1–10.
- [17] X. Chen, K. Gatsis, H. Hassani, and S. S. Bidokhti, "Age of information in random access channels," *IEEE Trans. Inf. Theory*, vol. 68, no. 10, pp. 6548–6568, Oct. 2022.
- [18] D. Ghose, F. Y. Li, and V. Pla, "MAC protocols for wake-up radio: Principles, modeling and performance analysis," *IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 2294–2306, May 2018.
- [19] A. Trotta et al., "BEE-DRONES: Energy-efficient data collection on wake-up radio-based wireless sensor networks," in *Proc. IEEE Conf. Comput. Commun. Workshops (IEEE INFOCOM)*, Apr. 2019, pp. 547–553.
- [20] J. Shiraishi, A. E. Kalør, F. Chiariotti, I. Leyva-Mayorga, P. Popovski, and H. Yomo, "Query timing analysis for content-based wake-up realizing informative IoT data collection," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 327–331, Feb. 2023.
- [21] J. Shiraishi, S. Cavallero, S. R. Pandey, F. Saggese, and P. Popovski, "Coexistence of push wireless access with pull communication for content-based wake-up radios," 2024, *arXiv:2404.12816*.
- [22] I. Kadota, A. Sinha, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Scheduling policies for minimizing age of information in broadcast wireless networks," *IEEE/ACM Trans. Netw.*, vol. 26, no. 6, pp. 2637–2650, Dec. 2018.
- [23] Y. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, vol. 19, no. 12, pp. 2903–2915, Dec. 2020.
- [24] S. S. Vilni, M. Moltafet, M. Leinonen, and M. Codreanu, "Multi-source AoI-constrained resource minimization under HARQ: Heterogeneous sampling processes," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 1084–1099, Jan. 2023.
- [25] N. Gast, "Optimization and control of large systems: Fighting the curse of dimensionality," Ph.D. thesis, Université Grenoble Alpes, Grenoble, France, 2010.
- [26] A. Maatouk, S. Kriouile, M. Assad, and A. Ephremides, "On the optimality of the Whittle's index policy for minimizing the age of information," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1263–1277, Feb. 2021.

- [27] S. Kriouile, M. Assaad, and A. Maatouk, "On the global optimality of Whittle's index policy for minimizing the age of information," *IEEE Trans. Inf. Theory*, vol. 68, no. 1, pp. 572–600, Jan. 2022.
- [28] C. Li, S. Li, Q. Liu, Y. T. Hou, W. Lou, and S. Kompella, "Eywa: A general approach for scheduler design in AoI optimization," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, May 2023, pp. 1–9.
- [29] E. O. Gamgam, N. Akar, and S. Ulukus, "Cyclic scheduling for age of information minimization with generate at will status updates," in *Proc. 58th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2024.
- [30] N. Akar, S. Liyanaarachchi, and S. Ulukus, "Scalable cyclic schedulers for age of information optimization in large-scale status update systems," in *Proc. IEEE Conf. Comput. Commun. Workshops (IEEE INFOCOM)*, May 2024, pp. 1–7.
- [31] R. D. Yates, "The age of information in networks: Moments, distributions, and sampling," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5712–5728, Sep. 2020.
- [32] M. Moltafet, M. Leinonen, and M. Codreanu, "Moment generating function of age of information in multisource M/G/1/1 queueing systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6503–6516, Oct. 2022.
- [33] N. Akar and E. O. Gamgam, "Distribution of age of information in status update systems with heterogeneous information sources: An absorbing Markov chain-based approach," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2024–2028, Aug. 2023.
- [34] A. M. Bedewy, Y. Sun, S. Kompella, and N. B. Shroff, "Optimal sampling and scheduling for timely status updates in multi-source networks," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4019–4034, Jun. 2021.
- [35] Y. Sun, E. Uysal-Biyikoglu, and S. Kompella, "Age-optimal updates of multiple information flows," in *Proc. IEEE Conf. Comput. Commun. Workshops (IEEE INFOCOM WKSHPS)*, Apr. 2018, pp. 136–141.
- [36] Y. Sun and S. Kompella, "Age-optimal multi-flow status updating with errors: A sample-path approach," *J. Commun. Netw.*, vol. 25, no. 5, pp. 570–584, Oct. 2023.
- [37] H. B. Beytur and E. Uysal-Biyikoglu, "Minimizing age of information for multiple flows," in *Proc. IEEE Int. Black Sea Conf. Commun. Netw. (BlackSeaCom)*, Jun. 2018, pp. 1–5.
- [38] B. Buyukates and S. Ulukus, "Timely communication in federated learning," in *Proc. IEEE Conf. Comput. Commun. Workshops (IEEE INFOCOM)*, May 2021, pp. 1–6.
- [39] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. V. Poor, "Age-based scheduling policy for federated learning in mobile edge networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 8743–8747.
- [40] F. Liu, W. S. Wong, Y.-H. Lo, Y. Zhang, C. S. Chen, and G. Xing, "Age of information for periodic status updates under sequence based scheduling," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 5963–5978, Oct. 2023.
- [41] C. Li, Q. Liu, S. Li, Y. Chen, Y. T. Hou, and W. Lou, "On scheduling with AoI violation tolerance," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, May 2021, pp. 1–9.
- [42] M. Shreedhar and G. Varghese, "Efficient fair queueing using deficit round Robin," in *Proc. ACM SIGCOMM*, Sep. 1995, pp. 231–242.

Sahan Liyanaarachchi (Graduate Student Member, IEEE) received the B.Sc. degree in electronic and telecommunication engineering from the University of Moratuwa, Sri Lanka, in 2020. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. His research interests include information theory, age of information, remote estimation, and machine learning.

Sennur Ulukus (Fellow, IEEE) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University and the Ph.D. degree in electrical and computer engineering from the Wireless Information Network Laboratory (WINLAB), Rutgers University.

She is a Distinguished University Professor and the Anthony Ephremides Professor in Information Sciences and Systems with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA, where she also holds a joint appointment with the Institute for Systems Research (ISR). Prior to joining UMD, she was a Senior Technical Staff Member with the AT&T Laboratories-Research. Her research interests include information theory, wireless communications, machine learning, signal processing, and networks; with a recent focus on the age of information, private information retrieval, blockchain consensus, quantum information, machine learning for wireless, federated learning, and large language models for communications and networks.

Dr. Ulukus is a Distinguished Scholar-Teacher of the University of Maryland. She received the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the 2019 IEEE Communications Society Best Tutorial Paper Award, the 2020 IEEE Communications Society Women in Communications Engineering (WICE) Outstanding Achievement Award, the 2020 IEEE Communications Society Technical Committee on Green Communications and Computing (TCGCC) Distinguished Technical Achievement Recognition Award, the 2024 IEEE Communications Society Communication Theory Technical Committee (CTTC) Technical Achievement Award, the 2005 NSF CAREER Award, the 2011 ISR Outstanding Systems Engineering Faculty Award, and the 2012 ECE George Corcoran Outstanding Teaching Award. She was a Distinguished Lecturer of the IEEE Information Theory Society (2018–2019). She is the TPC Chair of 2021 IEEE Globecom and was the TPC Co-Chair of 2019 IEEE ITW, 2017 IEEE ISIT, 2016 IEEE Globecom, 2014 IEEE PIMRC, and 2011 IEEE CTW. She was an Area Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2019–2023), an Area Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING (2016–2020), an Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS-Series on Green Communications and Networking (2015–2016), an Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY (2007–2010), and an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS (2003–2007). She was a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN INFORMATION THEORY (2021, 2022, and 2023), IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (2008, 2015, 2021, and 2022), IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (2021), *Journal of Communications and Networks* (2012), and IEEE TRANSACTIONS ON INFORMATION THEORY (2011). She is the TPC Chair of the 2021 IEEE Globecom and the TPC Co-Chair of the 2024 IEEE Globecom, the 2024 IEEE DySPAN, the 2023 IEEE MILCOM, the 2019 IEEE ITW, the 2017 IEEE ISIT, the 2016 IEEE Globecom, the 2014 IEEE PIMRC, and the 2011 IEEE CTW.

Nail Akar (Senior Member, IEEE) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Türkiye, in 1987, and the M.S. and Ph.D. degrees in electrical and electronics engineering from Bilkent University, Ankara, Türkiye, in 1989 and 1994, respectively.

From 1994 to 1996, he was a Visiting Scholar and a Visiting Assistant Professor with the Computer Science Telecommunications Program, University of Missouri-Kansas City, Kansas City, MO, USA. In 1996, he joined the Technology Planning and Integration Group, Long Distance Division, Sprint, Overland Park, Kansas, where he held a Senior Member of technical staff position from 1999 to 2000. Since 2000, he has been with Bilkent University, where he is currently a Professor with the Electrical and Electronics Engineering Department. He visited the School of Computing, University of Missouri-Kansas City, as a Fulbright Scholar in 2010. From 2023 to 2024 academic year, he was on sabbatical leave at the University of Maryland, as a Visiting Professor. His current research interests include information freshness, the age of information, green communications and networking, applications of queueing theory, and optimization in communication systems and blockchains. He is currently an Editor of IEEE TRANSACTIONS ON COMMUNICATIONS.