Timely Updates By Multiple Sources: The M/M/1 Queue Revisited

Sanjit K. Kaul
Wireless Systems Lab, IIIT-Delhi
email : skkaul@iiitd.ac.in

Roy D. Yates
WINLAB, ECE Dept., Rutgers University
email: ryates@winlab.rutgers.edu

2020 54th Annual Conference on Information Sciences and Systems (CISS) 978-1-7281-4085-8/20/$31.00 ©2020 IEEE 10.1109/CISS48834.2020.1570627741

***Abstract*—Multiple sources submit updates to a monitor through an M/M/1 queue. A stochastic hybrid system (SHS) approach is used to derive the average age of information (AoI) for an individual source as a function of the offered load of that source and the competing update traffic offered by other sources. This work corrects an error in a prior analysis. By numerical evaluation, this error is observed to be small and qualitatively insignificant.**

I. Introduction

Fueled by ubiquitous connectivity and advancements in portable devices, *real-time status* updates have become increasingly popular. This has driven emerging analytical interest in the Age of Information (AoI) metric for characterizing the timeliness of these updates [1]. Specifically, an update packet with time-stamp u is said to have age $t - u$ at a time $t \geqslant u$. When the monitor's freshest received update at time t has time-stamp $u(t)$, the age is the random sawtooth process $\Delta(t) = t - u(t)$ depicted in Figure 1. Optimization based on AoI metrics of both the network and the senders' updating policies has yielded new and even surprising results [2], [3].

Prior age analyses of the M/M/1 queue are based on graphical analysis of the sawtooth age process. As indicated by the title, this work revisits age analysis for updates from multiple sources passing through the M/M/1 queue using the stochastic hybrid systems (SHS) method. Prior use of SHS for age analysis has been restricted to systems in which old updates in the system are purged when new updates arrive. This resulted in a finite discrete state space, a technical requirement of SHS analysis. The challenge in this work is to apply SHS to an M/M/1 queue in which the discrete state, namely the queue backlog, can be arbitrarily large. Our approach is to employ SHS for age analysis of the M/M/1/m queue in which the system discards arriving updates that find m previously queued updates and then allow $m \to \infty$.

The main result, Theorem 2, can be shown to be numerically identical to the independently derived result in [4, Equation (55)]. In fact, both Theorem 2 and [4, Equation (55)] correct a faulty age analysis of the multi-source M/M/1 queue in [5] that propagated to [6].

A. Prior Work

AoI analysis of updating systems started with the analyses of status age in single-source single-server queues [1], the M/M/1 LCFS queue with preemption in service [7], and the M/M/1 FCFS system with multiple sources [5]. Since these initial efforts, there have been a large number of contributions to AoI analysis. This section summarizes work related to the analysis of age, with an emphasis on systems in which updates arrive as stochastic processes to queues and networks.

To evaluate AoI for a single source sending updates through a network cloud [8] or through an M/M/2 server [9], [10], out-of-order packet delivery was the key analytical challenge. A related (and generally more tractable) metric, peak age of information (PAoI), was introduced in [11]. Properties of PAoI were also studied for various M/M/1 queues that support preemption of updates in service or discarding of updates that find the server busy [12], [13]. In [11], [12], the authors analyzed AoI and PAoI for queues that discard arriving updates if the system is full and also for a third queue in which an arriving update would preempt a waiting update.

For a single updating source, distributional properties of the age process were analyzed for the D/G/1 queue under FCFS [14], as well as for single server FCFS and LCFS queues [15]. Packet deadlines were found to improve AoI [16]. Age-optimal preemption policies were identified for updates with deterministic service times [17]. AoI was evaluated in the presence of packet erasures at the M/M/1 queue output [18] and for memoryless arrivals to a two-state Markov-modulated service process [19].

There have also been efforts to evaluate and optimize age for multiple sources sharing a queue or simple network [6], [20]–[25]. In [6], the SHS approach was introduced to extend AoI results to preemptive queues with multiple sources. SHS has also been employed in [25] to evaluate a two-source M/M/1 queue in which the most recent update from each source is queued. With synchronized arrivals, a maximum age first policy was shown to be optimal under preemption in service and near-optimal for non-preemptive service [26]. A similar maximum age matching approach was analyzed for an orthogonal channel system [27]. Scheduling based on the Whittle index was also shown to perform nearly optimally [21], [28], [29]. AoI analysis of preemptive priority service systems has also been explored [30], [31]. Updates through communication channels have also been studied, including hybrid ARQ [32]–[36] for channels with bit erasures, and channel coding by an energy harvesting source [37].

The first evaluation of the average AoI over multihop network routes [38] employed a discrete-time version of the

978-1-7281-4085-8/20/$31.00 ©2020 IEEE

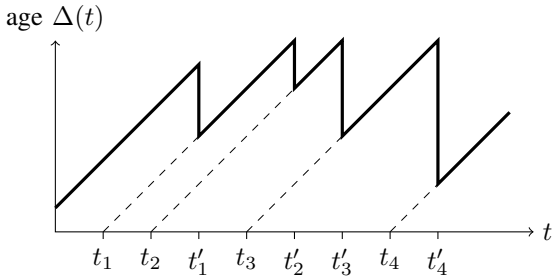

Fig. 1. Updates arrive at times t_k and are received by the monitor at times t'_k. At instants t'_k, the age is reset to $t'_k - t_k$, which is the time spent by the update in the system. Age increases linearly otherwise.

status sampling network subsequently introduced in [39]. When multiple sources employ wireless networks subject to interference constraints, AoI has been analyzed under a variety of link scheduling strategies [40]–[44]. When update transmission times are exponentially distributed, sample path arguments have shown that a preemptive Last-Generated, First-Served (LGFS) policy results in smaller age processes at all nodes of the network than any other causal policy [45]–[47].

B. System Model

Status updates from N independent sources are queued in a first-come-first-served (FCFS) manner at a single server facility. On completion of service, an update is delivered to a monitor, as illustrated in Figure 2. Updates of source i arrive as a Poisson process of rate λ_i and service times are exponentially distributed with rate μ. The load offered by user i is $\rho_i = \lambda_i/\mu$. The total offered load is $\rho = \sum_{i=1}^{N} \rho_i$ and the corresponding total rate of arrivals of updates is λ. For queue stability we require $\rho < 1$.

Each update is associated with a timestamp that records the time it was generated by its source. For this work, it is also the time the update arrives at the facility. At time t, let $u_i(t)$ be the timestamp of an update of source i most recently received by the monitor. The age of updates from source i at the monitor is the stochastic process $\Delta_i(t) = t - u_i(t)$. Figure 1 shows an illustration. We will calculate the average age $\Delta_i = \lim_{t\to\infty} \mathrm{E}[\Delta_i(t)]$. All sources other than i together constitute a Poisson process of rate $\lambda_{-i} = \sum_{j\neq i} \lambda_j$ and a corresponding offered load of $\rho_{-i} = \lambda_{-i}/\mu = \rho - \rho_i$.

II. Stochastic Hybrid Systems for AoI

As in [6], we model the system as a stochastic hybrid system (SHS) with hybrid state $(q(t), \mathbf{x}(t))$. Here, $q(t)$ is discrete and represents a Markov state sufficient for the characterization of age, while the row vector $\mathbf{x}(t) \in \mathbb{R}^{n+1}$ is continuous and captures the evolution of $n+1$ age-related processes.

In this work, we use the SHS described in [48] in a simplified form in which $\mathbf{x}(t)$ is a piecewise linear process. We now summarize this simplified SHS; further details can be found in [6] and references therein.

The discrete state $q(t) \in \mathcal{Q} = \{0, \ldots, m\}$ is the number of update packets in the system. In the graphical representation

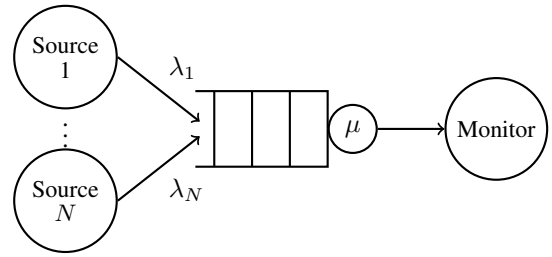

Fig. 2. N sources send updates through a shared queue to a monitor.

of the Markov chain $q(t)$, each state $q \in \mathcal{Q}$ is a node and each transition l is a directed edge (q_l, q'_l) with transition rate $\lambda^{(l)}\delta_{q_l,q(t)}$. Note that the Kronecker delta function $\delta_{q_l,q}$ ensures that transition l occurs only in state q_l. For each transition l, there is a transition reset mapping that can induce discontinuous jumps in the continuous state $\mathbf{x}(t)$. For AoI analysis, we employ a linear mapping of the form $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$. That is, transition l causes the system to jump to discrete state q'_l and resets the continuous state from $\mathbf{x}$ to $\mathbf{x}' = \mathbf{x}\mathbf{A}_l$. For tracking of the age process, the transition reset maps are binary: $\mathbf{A}_l \in \{0,1\}^{(n+1)\times(n+1)}$. Moreover, in each discrete state $q(t) = q$, the continuous state evolves as

$$\dot{\mathbf{x}}(t) = \mathbf{b}_q. \tag{1}$$

In using a piecewise linear SHS for AoI, the elements of $\mathbf{b}_q$ will be binary. We will see that the ones in $\mathbf{b}_q$ correspond to certain relevant components of $\mathbf{x}(t)$ that grow at unit rate in state q while the zeros mark components of $\mathbf{x}(t)$ that are irrelevant in state q to the age process and need not be tracked.

The transition rates $\{\lambda^{(l)}\}$ correspond to the transition rates associated with the continuous-time Markov chain for the discrete state $q(t)$; but there are some differences. Unlike an ordinary continuous-time Markov chain, the SHS may include self-transitions in which the discrete state is unchanged because a reset occurs in the continuous state. Furthermore, for a given pair of states $i, j \in \mathcal{Q}$, there may be multiple transitions l and l' in which the discrete state jumps from i to j but the transition maps $\mathbf{A}_l$ and $\mathbf{A}_{l'}$ are different.

It will be sufficient for average age analysis to define for all $\hat{q} \in \mathcal{Q}$,

$$\pi_{\hat{q}}(t) = \mathrm{E}\big[\delta_{\hat{q},q(t)}\big], \tag{2a}$$

$$v_{\hat{q}j}(t) = \mathrm{E}\big[x_j(t)\delta_{\hat{q},q(t)}\big], \quad 0 \leqslant j \leqslant n, \tag{2b}$$

and the vector functions

$$\mathbf{v}_{\hat{q}}(t) = [v_{\hat{q}0}(t), \ldots, v_{\hat{q}n}(t)] = \mathrm{E}\big[\mathbf{x}(t)\delta_{\hat{q},q(t)}\big]. \tag{2c}$$

Note that $\pi_{\hat{q}}(t) = \mathrm{E}\big[\delta_{\hat{q},q(t)}\big] = \mathrm{P}[q(t) = \hat{q}]$ denote the discrete state probabilities. Similarly, $\mathbf{v}_{\hat{q}}(t)$ is the conditional expectation of the age process, given that $q(t) = \hat{q}$, weighed by the probability of being in $\hat{q}$.

Let $\mathcal{L}$ be the set of all transitions. For each state q, let

$$\mathcal{L}'_q = \big\{l \in \mathcal{L} : q'_l = q\big\} \quad \text{and} \quad \mathcal{L}_q = \{l \in \mathcal{L} : q_l = q\} \tag{3}$$

denote the respective sets of incoming and outgoing transitions. A foundational assumption for age analysis is that the

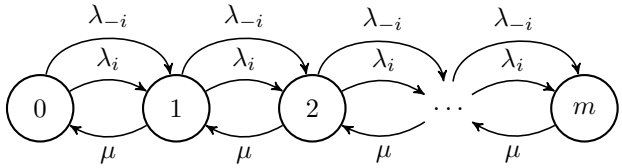

Fig. 3. The SHS Markov chain for transitions between discrete states.

Markov chain $q(t)$ is ergodic; otherwise, time-average age analysis makes little sense. Under this assumption, the state probability vector $\boldsymbol{\pi}(t) = \begin{bmatrix}\pi_0(t) & \cdots & \pi_m(t)\end{bmatrix}$ always converges to the unique stationary vector $\bar{\boldsymbol{\pi}} = \begin{bmatrix}\bar{\pi}_0 & \cdots & \bar{\pi}_m\end{bmatrix}$ satisfying

$$\bar{\pi}_{\bar{q}} \sum_{l \in \mathcal{L}_{\bar{q}}} \lambda^{(l)} = \sum_{l \in \mathcal{L}'_{\bar{q}}} \lambda^{(l)} \bar{\pi}_{q_l}, \quad \bar{q} \in \mathcal{Q}, \tag{4a}$$

$$\sum_{\bar{q} \in \mathcal{Q}} \bar{\pi}_{\bar{q}} = 1. \tag{4b}$$

When $\boldsymbol{\pi}(t) = \bar{\boldsymbol{\pi}}$, $\mathbf{v}(t) = \begin{bmatrix}\mathbf{v}_0(t) & \cdots & \mathbf{v}_m(t)\end{bmatrix}$ has been shown [6] to obey a system of first order differential equations, such that for all $\bar{q} \in \mathcal{Q}$,

$$\dot{\mathbf{v}}_{\bar{q}}(t) = \mathbf{b}_{\bar{q}} \bar{\pi}_{\bar{q}} + \sum_{l \in \mathcal{L}'_{\bar{q}}} \lambda^{(l)} \mathbf{v}_{q_l}(t) \mathbf{A}_l - \mathbf{v}_{\bar{q}}(t) \sum_{l \in \mathcal{L}_{\bar{q}}} \lambda^{(l)}. \tag{5}$$

Depending on the reset maps $\mathbf{A}_l$, the differential equation (5) may or may not be stable. However, when (5) is stable, each $\mathbf{v}_{\bar{q}}(t) = \mathrm{E}\big[\mathbf{x}(t)\delta_{\bar{q},q(t)}\big]$ converges to a limit $\bar{\mathbf{v}}_{\bar{q}}$ as $t \to \infty$. In this case, it follows that

$$\mathrm{E}[\mathbf{x}] = \lim_{t\to\infty} \mathrm{E}[\mathbf{x}(t)] = \lim_{t\to\infty} \sum_{\bar{q}\in\mathcal{Q}} \mathrm{E}\big[\mathbf{x}(t)\delta_{\bar{q},q(t)}\big] = \sum_{\bar{q}\in\mathcal{Q}} \bar{\mathbf{v}}_{\bar{q}}.$$

Here we follow the convention adopted in [6] that $x_0(t) = \Delta(t)$ is the age at the monitor. With this convention, the average age of the process of interest is then $\Delta = \mathrm{E}[x_0] = \sum_{\bar{q}\in\mathcal{Q}} \bar{v}_{\bar{q}0}$. The following theorem provides a simple way to calculate the average age in an ergodic queueing system.

Theorem 1. *[6, Theorem 4] If the discrete-state Markov chain $q(t)$ is ergodic with stationary distribution $\bar{\boldsymbol{\pi}}$ and we can find a non-negative solution $\bar{\mathbf{v}} = \begin{bmatrix}\bar{\mathbf{v}}_0 & \cdots \bar{\mathbf{v}}_m\end{bmatrix}$ such that*

$$\bar{\mathbf{v}}_{\bar{q}} \sum_{l \in \mathcal{L}_{\bar{q}}} \lambda^{(l)} = \mathbf{b}_{\bar{q}} \bar{\pi}_{\bar{q}} + \sum_{l \in \mathcal{L}'_{\bar{q}}} \lambda^{(l)} \bar{\mathbf{v}}_{q_l} \mathbf{A}_l, \qquad \bar{q} \in \mathcal{Q}, \tag{6}$$

then the differential equation (5) is stable and the average age of the AoI SHS is given by $\Delta = \sum_{\bar{q}\in\mathcal{Q}} \bar{v}_{\bar{q}0}$.

III. SHS Analysis of Age

The discrete Markov state $q(t) \in \{0, 1, 2, \ldots\}$ tracks the number of updates in the system at time t. Figure 3 shows the evolution of the state for a corresponding blocking system that has a finite occupancy of m. In states $0 < k < m$, an arrival, which results from a rate λ transition, increases the state to $k+1$. A departure on completion of service, which results from a rate μ transition, reduces the occupancy to $k-1$. An empty system, $q(t) = 0$, can only see an arrival; while when the system is full, $q(t) = m$, all arrivals are blocked and cleared and only a departure may take place.

TABLE I
SSH transitions and reset maps for the chain in Figure 3.

$q_l \to q'_l$	$\lambda^{(l)}$	$\mathbf{x}\mathbf{A}_l$	$\mathbf{A}_l$	$\mathbf{v}_{q_l}\mathbf{A}_l$
$1 \to 0$	μ	$\begin{bmatrix}x_1 & \mathbf{0}\end{bmatrix}$	$\mathbf{D}_0$	$\begin{bmatrix}v_{1,1} & \mathbf{0}\end{bmatrix}$
$0 \to 1$	λ_i	$\begin{bmatrix}x_0 & 0 & \mathbf{0}\end{bmatrix}$	Λ_1	$\begin{bmatrix}v_{0,0} & 0 & \mathbf{0}\end{bmatrix}$
$0 \to 1$	λ_{-i}	$\begin{bmatrix}x_0 & x_0 & \mathbf{0}\end{bmatrix}$	Γ_1	$\begin{bmatrix}v_{0,0} & v_{0,0} & \mathbf{0}\end{bmatrix}$
$2 \to 1$	μ	$\begin{bmatrix}x_1 & x_2 & \mathbf{0}\end{bmatrix}$	$\mathbf{D}_1$	$\begin{bmatrix}v_{21} & v_{22} & \mathbf{0}\end{bmatrix}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$k-1 \to k$	λ_i	$\begin{bmatrix}(\mathbf{x})_{0:k-1} & 0 & \mathbf{0}\end{bmatrix}$	Λ_k	$\begin{bmatrix}(\mathbf{v}_{k-1})_{0:k-1} & 0 & \mathbf{0}\end{bmatrix}$
$k-1 \to k$	λ_{-i}	$\begin{bmatrix}(\mathbf{x})_{0:k-1} & x_{k-1} & \mathbf{0}\end{bmatrix}$	Γ_k	$\begin{bmatrix}(\mathbf{v}_{k-1})_{0:k-1} & v_{k-1,k-1} & \mathbf{0}\end{bmatrix}$
$k+1 \to k$	μ	$\begin{bmatrix}(\mathbf{x})_{1:k+1} & \mathbf{0}\end{bmatrix}$	$\mathbf{D}_k$	$\begin{bmatrix}(\mathbf{v}_{k+1})_{1:k+1} & \mathbf{0}\end{bmatrix}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$m-1 \to m$	λ_i	$\begin{bmatrix}(\mathbf{x})_{0:m-1} & 0 & 0\end{bmatrix}$	Λ_m	$\begin{bmatrix}(\mathbf{v}_{m-1})_{0:m-1} & 0 & 0\end{bmatrix}$
$m-1 \to m$	λ_{-i}	$\begin{bmatrix}(\mathbf{x})_{0:m-1} & x_{m-1} & 0\end{bmatrix}$	Γ_m	$\begin{bmatrix}(\mathbf{v}_{m-1})_{0:m-1} & v_{m-1,m-1} & 0\end{bmatrix}$

Let Δ_i^m be the average age of source i for the size m blocking system. The average age Δ_i of source i at the monitor can be obtained as

$$\Delta_i = \lim_{m\to\infty} \Delta_i^m. \tag{7}$$

For the blocking system, $\mathbf{x}(t) = \begin{bmatrix}x_0(t) & x_1(t) & \ldots & x_m(t)\end{bmatrix}$ is the age state. Here $x_0(t)$ is the age process of source i at the monitor. The age process $x_j(t)$, $1 \leqslant j \leqslant k$, tracks the age to which the age $x_0(t)$ will be reset when the packet currently in the jth position in the queue completes service, where $j = 1$ corresponds to the packet in service. Note that since an update packet in the queue may not be of source i, $x_j(t)$ is in general not the age of the update itself. Also, in any discrete state k we must track only the $k+1$ age processes $x_0(t), x_1(t), \ldots, x_k(t)$. The rest are irrelevant. We assume that in state $q(t) = k$, $x_j(t) = 0$ for $j > k$. The age state evolves according to

$$\dot{\mathbf{x}}(t) = \mathbf{b}_k \equiv \begin{bmatrix}\mathbf{1}_{k+1} & \mathbf{0}_{m-k}\end{bmatrix}. \tag{8}$$

We use the shorthand notation $(\mathbf{x})_{j:l}$ and $(\mathbf{v}_q)_{j:l}$, respectively, to denote the vectors $\begin{bmatrix}x_j & \cdots & x_l\end{bmatrix}$ and $\begin{bmatrix}v_{q,j} & \ldots & v_{q,l}\end{bmatrix}$. Table I shows the transition reset maps for the transitions in the Markov chain in Figure 3. We also show the age state $\mathbf{x}_{q'}$ and the vector $\mathbf{v}_{q'}$ obtained on transitioning from state q to q'.

Consider an arrival that transitions the system into state $0 < k < m$ from $k-1$ as a result of an arrival from source i. The age processes $x_0(t), x_1(t), \ldots, x_{k-1}(t)$ stay unaffected by the transition and continue to evolve as in $k-1$. That is $x'_j = x_j$, $0 \leqslant j \leqslant k-1$. In state k we must in addition track $x_k(t)$. We set $x_k(t) = 0$, since the age of this new update from source i is 0. Further, $\dot{x}_k(t) = 1$. The transition reset map is given by

$$\Lambda_k = \left[\begin{array}{c|c} \underbrace{\begin{matrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{matrix}}_{k\times(k+1)} & \mathbf{0} \\ \mathbf{0} & \end{array}\right]. \tag{9}$$

If the arrival is of another source, we must set $x_k(t)$ in a

manner such that when this arrival completes service, it must not change the age of i at the monitor. We set $x_k(t) = x_{k-1}(t)$. As before, $\dot{x}_k(t) = 1$. Note that in this case we are effectively setting $x_k(t)$ to the age of the freshest update of source i, if any, in the system. If there is no update of source i in the system, we are effectively setting $x_k(t)$ to $x_0(t)$ as in this case all relevant age processes $x_0(t), x_1(t), \ldots, x_k(t)$ will be tracking the age at the monitor. The transition reset map is

$$\Gamma_k = \left[\begin{array}{c|c} \underbrace{\begin{matrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \end{matrix}}_{k\times(k+1)} & \mathbf{0} \\ \mathbf{0} & \end{array}\right]. \tag{10}$$

Now consider a departure in state $k+1$. The SSH transitions to state k. The age process $x_1(t)$ in $k+1$ corresponds to the departure. On completion of service, this resets the process $x_0(t)$ at the monitor in state k. Once the packet departs, the one next in queue enters service. Thus the process $x_2(t)$ in state $k+1$ resets $x_1(t)$ in state k and so on. The reset map is given by

$$\mathbf{D}_k = \left[\begin{array}{c|c} \underbrace{\begin{matrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{matrix}}_{(k+2)\times(k+1)} & \mathbf{0} \\ \mathbf{0} & \end{array}\right]. \tag{11}$$

We can now write the equations given by (6) in Theorem 1 for our blocking system.

$$\lambda\bar{\mathbf{v}}_0 = \mathbf{b}_0\bar{\pi}_0 + \mu\bar{\mathbf{v}}_1\mathbf{D}_0, \tag{12a}$$

$$\begin{aligned}(\lambda+\mu)\bar{\mathbf{v}}_k &= \mathbf{b}_k\bar{\pi}_k + \lambda_i\bar{\mathbf{v}}_{k-1}\Lambda_k \\ &\quad + \lambda_{-i}\bar{\mathbf{v}}_{k-1}\Gamma_k + \mu\bar{\mathbf{v}}_{k+1}\mathbf{D}_k,\ 0<k<m,\end{aligned} \tag{12b}$$

$$\mu\bar{\mathbf{v}}_m = \mathbf{b}_m\bar{\pi}_m + \lambda_i\bar{\mathbf{v}}_{m-1}\Lambda_m + \lambda_{-i}\bar{\mathbf{v}}_{m-1}\Gamma_m. \tag{12c}$$

We extract equations corresponding to the discrete states $\bar{q}$ and the relevant age processes $x_j(t)$ in the states. We have $k+1$ relevant age processes in state $0<k<m$. For such a state k and $x_j(t)$, $0 \leqslant j \leqslant k$, we have

$$\begin{aligned}(\lambda+\mu)\bar{v}_{k,j} &= \bar{\pi}_k + \lambda\bar{v}_{k-1,j} \\ &\quad + \mu\bar{v}_{k+1,j+1}, \quad 0\leqslant j\leqslant k-1,\end{aligned} \tag{13a}$$

$$(\lambda+\mu)\bar{v}_{k,k} = \bar{\pi}_k + \lambda_{-i}\bar{v}_{k-1,k-1} + \mu\bar{v}_{k+1,k+1}. \tag{13b}$$

In state $k=0$, no departures take place and the only relevant age process is $x_0(t)$. We have

$$\lambda\bar{v}_{0,0} = \bar{\pi}_0 + \mu\bar{v}_{1,1}. \tag{14}$$

Lastly, in state $k=M$, arrivals are inconsequential while a departure may take place. All age processes $x_0(t), \ldots, x_m(t)$ are relevant. We have

$$\mu\bar{v}_{m,j} = \bar{\pi}_m + \lambda\bar{v}_{m-1,j},\ 0\leqslant j\leqslant n-1, \tag{15a}$$

$$\mu\bar{v}_{m,m} = \bar{\pi}_m + \lambda_{-i}\bar{v}_{m-1,m-1}. \tag{15b}$$

The steady state probability $\bar{\pi}_k$ of k updates in the blocking M/M/1 system can be obtained by solving (4a)-(4b). The equations are

$$\lambda\bar{\pi}_0 = \mu\bar{\pi}_1, \tag{16a}$$

$$(\lambda+\mu)\bar{\pi}_k = \lambda\bar{\pi}_{k-1} + \mu\bar{\pi}_{k+1},\ 0<k<m, \tag{16b}$$

$$\mu\bar{\pi}_m = \lambda\bar{\pi}_{m-1}. \tag{16c}$$

Combined with the normalization constraint $\sum_{q=0}^{m}\bar{\pi}_q = 1$, (16) yields the M/M/1/$m$ stationary distribution

$$\bar{\pi}_k = \left(\frac{1-\rho}{1-\rho^{m+1}}\right)\rho^k, \quad 0\leqslant k\leqslant m. \tag{17}$$

From Theorem 1 we know that the average age of source i in the blocking system is

$$\Delta_i^m = \sum_{\bar{q}=0}^{m}\bar{v}_{\bar{q}0}. \tag{18}$$

The average age Δ_i is obtained using Equation (7). Details regarding calculation of Δ_i^m and Δ_i are in the Appendix. With the definition

$$\mathcal{E}_i \equiv \frac{1+\rho-\sqrt{(1+\rho)^2-4\rho_{-i}}}{2\rho_{-i}}, \tag{19}$$

we state our main result.

Theorem 2. *Poisson sources $1, 2, \ldots, N$ send their status updates to a monitor via an M/M/1 FCFS queue with service rate μ. The sources offer loads $\rho_1, \ldots, \rho_N$. The average age of updates of source i at the monitor is*

$$\Delta_i = \frac{1}{\mu}\left[\frac{1-\rho}{(\rho-\rho_{-i}\mathcal{E}_i)(1-\rho\mathcal{E}_i)} + \frac{1}{1-\rho} + \frac{\rho_{-i}}{\rho_i}\right]. \tag{20}$$

Note that the expression approaches the age of a single user M/M/1 queue as $\rho_{-i}\to 0$. This is easy to see on noting the fact that $\lim_{\rho_{-i}\to 0}\mathcal{E}_i = 1/(1+\rho)$.

IV. Evaluation

Figure 4 compares average age obtained using (20) with that estimated using simulation experiments. We plot the age Δ_1 of source 1 as a function of its offered load ρ_1 for various choices of the load ρ_2 offered by the other users. Each empirically estimated value of age was obtained by simulating on the order of 10^8 transitions of the Markov chain. As is shown in the figure, empirical results match our analysis. We also show the age that was obtained via analysis in our earlier work [6, Theorem 1]. We observe that the error in the age increases with the load of the other user. This error does disappear as $\rho_2 \to 0$. This is understandable as the error was introduced in accounting for the number of packets of other users in the system that are found by such an arrival of source 1 that arrives after the previous arrival of the source has completed service.

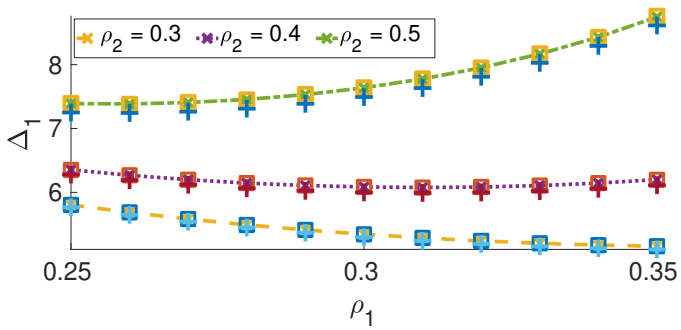

Fig. 4. Age of source 1 as a function of its utilization is shown for different values of utilization ρ_2 of the other source. For each ρ_2, the $\times$ markers correspond to the age given by (20), the $\square$ markers show the age obtained using simulations, and the $+$ show the age obtained in the earlier work that analyzed multiple memoryless sources [6, Theorem 1]. While (20) matches well with simulations, the $+$ drift away as ρ_2 increases. We set $\mu = 1$.

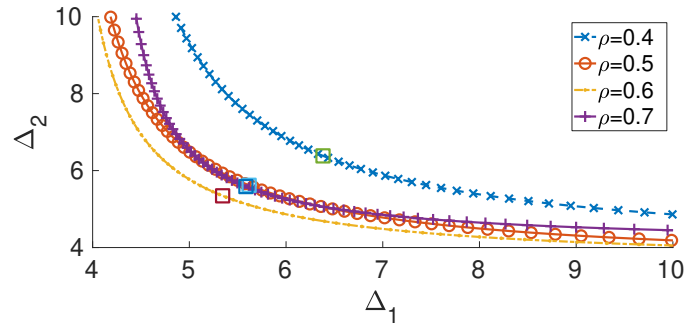

Fig. 5. Age contours obtained by varying the share of two sources of a given total utilization ρ. The squares $\square$ correspond to the ages obtained when each source offers a load of $\rho/2$. Sum age is minimized when $\rho \approx 0.6$ for $\rho_1 = \rho_2 \approx 0.3$.

Figure 5 shows average ages Δ_1 and Δ_2 of the sources as their relative loads are varied, for a fixed total load of ρ. The sum age is minimized when $\rho \approx 0.6$ and the sources share the load equally. At other relative loadings of the sources, the sum age may not be minimized at $\rho = 0.6$.

V. Conclusion

This work revisits average age analysis of status updates from multiple sources passing through the M/M/1 queue. SHS analysis of the M/M/1/m blocking queue enables a relatively simple analysis and closed form result for the average age.

Acknowledgement

We thank the authors of [4] for bringing the error in the M/M/1 analysis of [6] to our attention.

Appendix: Proof of Theorem 2

Proof. We start by calculating the average Δ_i^m of the blocking system as given in (18). We proceed by writing $\sum_{k=1}^m \bar{v}_{k,0}$ in terms of the probabilities $\bar{\pi}_k$, $0 \leqslant k \leqslant m$. That leaves us with the tedium of calculating $\bar{v}_{0,0}$. Using (13a), we can write

$$\lambda \bar{v}_{0,0} = \bar{\pi}_0 + \mu \bar{v}_{1,1}, \tag{21a}$$

$$\lambda \bar{v}_{1,0} + \mu \bar{v}_{1,0} - \lambda \bar{v}_{0,0} = \bar{\pi}_1 + \mu \bar{v}_{21}, \tag{21b}$$

$$\vdots$$

$$\lambda \bar{v}_{m-1,0} + \mu \bar{v}_{m-1,0} - \lambda \bar{v}_{m-2,0} = \bar{\pi}_{m-2} + \mu \bar{v}_{m,1}, \tag{21c}$$

$$\mu \bar{v}_{m,0} - \lambda \bar{v}_{m-1,0} = \bar{\pi}_m. \tag{21d}$$

Adding these equations we obtain

$$\sum_{k=1}^m \bar{v}_{k,0} = \frac{1}{\mu} \sum_{k=0}^m \bar{\pi}_k + \sum_{k=1}^m \bar{v}_{k,1}. \tag{22}$$

Similarly, using (13a) to write equations corresponding to $\bar{v}_{k,1}$, $1 \leqslant k \leqslant m$, we obtain

$$\sum_{k=1}^m \bar{v}_{k,1} = \frac{1}{\mu} \sum_{k=1}^m \bar{\pi}_k + \frac{\lambda_{-i}}{\mu} + \sum_{k=2}^m \bar{v}_{k,2}. \tag{23}$$

Proceeding similarly, we obtain expressions for $\sum_{k=2}^m \bar{v}_{k,2}, \sum_{k=3}^m \bar{v}_{k,3}, \ldots, \sum_{k=m-1}^m \bar{v}_{k,m-1}$. Further (15b) gives us $\bar{v}_{m,m}$. We can now rewrite (22) as

$$\sum_{k=1}^m \bar{v}_{k,0} = \frac{1}{\mu} \sum_{k=0}^m (k+1)\bar{\pi}_k + \frac{\lambda_{-i}}{\mu} \sum_{k=0}^{m-1} \bar{v}_{k,k}. \tag{24}$$

Now consider equations (13b) for $1 < k < m$, (14), and (15b). Summing them gives

$$\sum_{k=0}^{m-1} \bar{v}_{k,k} = \frac{1}{\lambda_i}. \tag{25}$$

Substituting in (24) allows us to write $\sum_{k=1}^m \bar{v}_{k,0}$ in terms of the steady state probabilities given by (17). To calculate $\bar{v}_{0,0}$, we will express $\bar{v}_{k,k}$ in terms of $\bar{v}_{0,0}$ and then use (25).

Rearranging terms in Equation (13b) we get

$$\bar{v}_{k,k} = \left(\frac{\lambda+\mu}{\mu}\right) \bar{v}_{k-1,k-1} - \left(\frac{\lambda_{-i}}{\mu}\right) \bar{v}_{k-2,k-2} - \frac{\bar{\pi}_{k-1}}{\mu}. \tag{26}$$

Repeated application of the equation to expand the terms $\bar{v}_{k-1,k-1}$ and $\bar{v}_{k-2,k-2}$ gives

$$\bar{v}_{k,k} = \left(c_{1,k}\frac{\lambda}{\mu} - c_{2,k}\frac{\lambda_{-i}}{\mu}\right) \bar{v}_{0,0} - \sum_{j=1}^k c_{j,k} \frac{\bar{\pi}_{j-1}}{\mu}. \tag{27}$$

The coefficients $c_{j,k}$ are obtained by solving the equations

$$c_{k-j,k} = \frac{\lambda+\mu}{\mu} c_{k-j+1,k} - \frac{\lambda_{-i}}{\mu} c_{k-j+2,k}, \tag{28}$$

with $c_{k,k} = 1$ and $0 \leqslant j \leqslant k$. We get

$$c_{j,k} = \frac{\left(\beta - \rho_{-i}^{-1}(1+\rho)\right)\left((1+\rho)\rho_{-i}^{-1} + \beta\right)^{j-k}}{2\beta} + \frac{\left(\beta + \rho_{-i}^{-1}(1+\rho)\right)\left((1+\rho)\rho_{-i}^{-1} - \beta\right)^{j-k}}{2\beta}, \tag{29}$$

where $\beta = \sqrt{(1+\rho)^2\rho_{-i}^{-2} - 4\rho_{-i}^{-1}}$. Substituting $c_{j,k}$ obtained above in (27) and substituting the resulting $\bar{v}_{k,k}$ in (25) gives us $\bar{v}_{0,0}$. Together with (24) and (25), we can obtain the average age Δ_i^m of the blocking system. Taking the limit as $m \to \infty$ gives us the average age Δ_i in Theorem 2. $\square$

REFERENCES

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

[2] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Info. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

[3] R. Yates, "Lazy is timely: Status updates by an energy harvesting source," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2015, pp. 3008–3012.

[4] M. Moltafet, M. Leinonen, and M. Codreanu, "On the age of information in multi-source queueing models," 2019. [Online]. Available: https://arxiv.org/abs/1911.07029

[5] R. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jul. 2012.

[6] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, 2018.

[7] S. Kaul, R. Yates, and M. Gruteser, "Status updates through queues," in *Conf. on Information Sciences and Systems (CISS)*, Mar. 2012.

[8] C. Kam, S. Kompella, and A. Ephremides, "Age of information under random updates," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, 2013, pp. 66–70.

[9] ——, "Effect of message transmission diversity on status age," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2014, pp. 2411–2415.

[10] C. Kam, S. Kompella, G. D. Nguyen, and A. Ephremides, "Effect of message transmission path diversity on status age," *IEEE Trans. Info. Theory*, vol. 62, no. 3, pp. 1360–1374, Mar. 2016.

[11] M. Costa, M. Codreanu, and A. Ephremides, "Age of information with packet management," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2014, pp. 1583–1587.

[12] ——, "On the age of information in status update systems with packet management," *IEEE Trans. Info. Theory*, vol. 62, no. 4, pp. 1897–1910, April 2016.

[13] V. Kavitha, E. Altman, and I. Saha, "Controlling packet drops to improve freshness of information," *CoRR*, vol. abs/1807.09325, 2018. [Online]. Available: http://arxiv.org/abs/1807.09325

[14] J. P. Champati, H. Al-Zubaidy, and J. Gross, "Statistical guarantee optimization for age of information for the D/G/1 queue," in *IEEE Conference on Computer Communications (INFOCOM) Workshops*, April 2018, pp. 130–135.

[15] Y. Inoue, H. Masuyama, T. Takine, and T. Tanaka, "A general formula for the stationary distribution of the age of information and its application to single-server queues," *CoRR*, vol. abs/1804.06139, 2018. [Online]. Available: http://arxiv.org/abs/1804.06139

[16] C. Kam, S. Kompella, G. D. Nguyen, J. Wieselthier, and A. Ephremides, "Age of information with a packet deadline," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, 2016, pp. 2564–2568.

[17] B. Wang, S. Feng, and J. Yang, "To skip or to switch? Minimizing age of information under link capacity constraint," *CoRR*, vol. abs/1806.08698, 2018. [Online]. Available: http://arxiv.org/abs/1806.08698

[18] K. Chen and L. Huang, "Age-of-information in the presence of error," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, 2016, pp. 2579–2584.

[19] L. Huang and L. P. Qian, "Age of information for transmissions over Markov channels," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2017.

[20] L. Huang and E. Modiano, "Optimizing age-of-information in a multi-class queueing system," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jun. 2015.

[21] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2016, pp. 844–851.

[22] S. K. Kaul and R. Yates, "Status updates over unreliable multiaccess channels," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jun. 2017, pp. 331–335.

[23] E. Najm and E. Telatar, "Status updates in a multi-stream M/G/1/1 preemptive queue," in *IEEE Conference on Computer Communications (INFOCOM) Workshops*, April 2018, pp. 124–129.

[24] Z. Jiang, B. Krishnamachari, X. Zheng, S. Zhou, and Z. Miu, "Decentralized status update for age-of-information optimization in wireless multiaccess channels," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2018, pp. 2276–2280.

[25] M. Moltafet, M. Leinonen, and M. Codreanu, "Average age of information for a multi-source m/m/1 queueing model with packet management," *arXiv preprint arXiv:2001.03959*, 2020.

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

[27] V. Tripathi and S. Moharir, "Age of information in multi-source systems," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2017.

[28] Z. Jiang, B. Krishnamachari, S. Zhou, and Z. Niu, "Can decentralized status update achieve universally near-optimal age-of-information in wireless multiaccess channels?" in *International Teletraffic Congress ITC 30*, September 2018. [Online]. Available: http://arxiv.org/abs/1803.08189

[29] Y.-P. Hsu, "Age of information: Whittle index for scheduling stochastic arrivals," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2018, pp. 2634–2638.

[30] E. Najm, R. Nasser, and E. Telatar, "Content based status updates," *CoRR*, vol. abs/1801.04067, 2018. [Online]. Available: http://arxiv.org/abs/1801.04067

[31] S. Kaul and R. Yates, "Age of information: Updates with priority," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jun. 2018, pp. 2644–2648.

[32] P. Parag, A. Taghavi, and J. Chamberland, "On real-time status updates over symbol erasure channels," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017.

[33] E. Najm, R. Yates, and E. Soljanin, "Status updates through M/G/1/1 queues with HARQ," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jun. 2017, pp. 131–135.

[34] R. Yates, E. Najm, E. Soljanin, and J. Zhong, "Timely updates over an erasure channel," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, Jun. 2017, pp. 316–320.

[35] E. T. Ceran, D. Gündüz, and A. György, "Average age of information with hybrid ARQ under a resource constraint," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, April 2018.

[36] H. Sac, T. Bacinoglu, E. Uysal-Biyikoglu, and G. Durisi, "Age-optimal channel coding blocklength for an M/G/1 queue with HARQ," in *19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2018, pp. 486–490.

[37] A. Baknina and S. Ulukus, "Coded status updates in an energy harvesting erasure channel," *CoRR*, vol. abs/1802.00431, 2018. [Online]. Available: http://arxiv.org/abs/1802.00431

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

[39] R. D. Yates, "The age of information in networks: Moments, distributions, and sampling," *arXiv preprint arXiv:1806.03487*, vol. abs/1806.03487, 2018.

[40] Q. He, D. Yuan, and A. Ephremides, "Optimal link scheduling for age minimization in wireless systems," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 5381–5394, July 2018.

[41] N. Lu, B. Ji, and B. Li, "Age-based scheduling: Improving data freshness for wireless real-time traffic," in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. Mobihoc '18. New York, NY, USA: ACM, 2018, pp. 191–200. [Online]. Available: http://doi.acm.org/10.1145/3209582.3209602

[42] R. Talak, S. Karaman, and E. Modiano, "Distributed scheduling algorithms for optimizing information freshness in wireless networks," *CoRR*, vol. abs/1803.06469, 2018. [Online]. Available: http://arxiv.org/abs/1803.06469

[43] ——, "Optimizing age of information in wireless networks with perfect channel state information," *CoRR*, vol. abs/1803.06471, 2018. [Online]. Available: http://arxiv.org/abs/1803.06471

[44] R. Talak, I. Kadota, S. Karaman, and E. Modiano, "Scheduling policies for age minimization in wireless networks with unknown channel state," *CoRR*, vol. abs/1805.06752, 2018. [Online]. Available: http://arxiv.org/abs/1805.06752

[45] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Optimizing data freshness, throughput, and delay in multi-server information-update systems," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, 2016, pp. 2569–2574.

[46] ——, "Age-optimal information updates in multihop networks," in *Proc. IEEE Int'l. Symp. Info. Theory (ISIT)*, June 2017, pp. 576–580.

[47] ——, "The age of information in multihop networks," *CoRR*, vol. abs/1712.10061, 2017. [Online]. Available: http://arxiv.org/abs/1712.10061

[48] J. Hespanha, "Modelling and analysis of stochastic hybrid systems," *IEE Proceedings-Control Theory and Applications*, vol. 153, no. 5, pp. 520–535, 2006.