Age of Information in G/G/1/1 Systems: Age Expressions, Bounds, Special Cases, and Optimization

Alkan Soysal, *Member, IEEE*, and Sennur Ulukus, *Fellow, IEEE*

***Abstract*— We consider the average age of information in G/G/1/1 systems under two service discipline models. In the first model, if a new update arrives when the service is busy, it is blocked; in the second model, a new update preempts the current update in service. For the blocking model, we first derive an exact age expression for G/G/1/1 systems. Then, using the age expression for G/G/1/1 systems, we calculate average age expressions for special cases, i.e., M/G/1/1 and G/M/1/1 systems. We observe that deterministic interarrivals minimize the average age of G/M/1/1 systems for a given mean interarrival time. Next, for the preemption in service model, we first derive an exact average age expression for G/G/1/1 systems. Then, similar to blocking discipline, using the age expression for G/G/1/1 systems, we calculate average age expressions for special cases, i.e., M/G/1/1 and G/M/1/1 systems. Average age for G/M/1/1 can be written as a summation of two terms, the first of which depends only on the first and second moments of interarrival times and the second of which depends only on the service rate. In other words, interarrival and service times are decoupled. We prove that deterministic interarrivals are optimum for G/M/1/1 systems for a given mean interarrival time. On the other hand, we observe for non-exponential service times that the optimal distribution of interarrival times depends on the relative values of the mean interarrival time and the mean service time. Finally, we propose a simple to calculate upper bound to the average age for the preemption in service discipline.**

***Index Terms*— Age of information, G/G/1/1 systems, blocking discipline, preemption in service discipline.**

I. Introduction

No matter how important information might be, there is a duration of time after which information loses its freshness. Especially in today's immensely interactive world, information ages fast. Hence, in recent years, researchers have begun to consider the age of information (AoI). Age of a status update, $\Delta(t) = t - u(t)$, is a random process which is defined as the difference between current time, t, and the time-stamped value of the most recently received update, $u(t)$. The AoI is defined as the average $\Delta(t)$ and given as [2]

$$\Delta = \lim_{T\to\infty} \frac{1}{T}\int_0^T \Delta(t)dt. \tag{1}$$

This definition is sufficiently broad to cover almost all communication scenarios. However, most of the AoI literature so far has considered queueing systems with Poisson arrival processes, mostly due to their mathematical tractability. On the other hand, Poisson arrivals (hence exponential interarrival times) might not correctly model the periodic and asynchronous nature of packet arrivals in certain wireless network applications [3], [4]. In this paper, we analyze AoI for general queueing distributions that might be observed in real world communication scenarios.

The first papers that consider the AoI in a communication setting are [5], [6], and [7]. The authors in [5] assume First Come First Served (FCFS) systems and calculates the average age expressions for M/M/1, M/D/1 and D/M/1 queues, [6] assumes Last Come First Served (LCFS) systems with and without preemption and calculates the average age expression for M/M/1 queues, [7] assumes multi-source FCFS systems with M/M/1 queues, and [2] provides a more detailed analysis. Starting with these works, there has been a growing interest in AoI analysis. For example, [8] considers a packet management approach for M/M/1/1 and M/M/1/2 queues, [9] calculates the average age for exponential interarrival and gamma service times, and [10] calculates the average age for an M/G/1/1 queue and finds the optimum arrival rate to minimize age. For exponential service times, preemption in service discipline is found to be optimal in [11] over all service disciplines.

Although they derive age expressions and propose age-minimum strategies, the main facilitator in the analyses of [2], [5]–[11] is the memoryless property of exponential interarrival times and/or exponential service times. Since arrival processes are not necessarily Poisson in a real world wireless network, some of these results do not hold in general. In this paper, we show examples of previous results that do not generalize to arbitrary interarrival and service time distributions. Such a general approach is also taken in [12], [13], where the focus is on the distribution of age rather than the average age that we consider in this paper. On the other

Manuscript received May 31, 2019; revised March 19, 2021; accepted June 14, 2021. Date of publication July 9, 2021; date of current version October 20, 2021. An earlier version of this paper was presented in part at the 2019 Asilomar Conference [1]. *(Corresponding author: Sennur Ulukus.)*

Alkan Soysal was with the Department of Electrical and Electronics Engineering, Bahçeşehir University, 34349 Istanbul, Turkey. He is now with the Calhoun Discovery Program, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA 24061 USA (e-mail: soysal@vt.edu).

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

Communicated by P. Mitran, Associate Editor for Communications.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIT.2021.3095823.

Digital Object Identifier 10.1109/TIT.2021.3095823

hand, [14] calculates the peak and average age of FCFS and LCFS G/G/1 queues with preemption in service discipline. In the case of LCFS with preemption in service discipline, the main result is in the form of a lower bound to the average age that does not reflect the effect of service time distribution.

While the literature on calculating age expressions for different queueing models expands, another line of research considers age minimization for energy constrained systems, i.e., energy harvesting problems. The goal is to find the optimum update generation policy that minimizes age, given the service time distribution and an energy constraint. In [15], the authors show the existence of an optimal stationary deterministic update generation policy when the service time process is a stationary and ergodic Markov chain. Age minimization for offline energy harvesting problems is considered in [16]–[18], and age minimization for online energy harvesting problems is considered in [19]–[22]. A recent survey paper provides a very detailed literature review on AoI [23].

The derivation of AoI for a given queue model requires a probabilistic approach in order to calculate the expected values of several, possibly correlated, random quantities specific to that model. The more complex the system is, the harder it is to calculate the expected values, especially when the interarrival times are not exponential. To overcome this, [2] proposes an approach based on stochastic hybrid systems (SHS). In this paper, we follow an alternative path to SHS in order to analyze general cases.

In this paper, our goal is to analyze AoI for general interarrival and service time distributions. Using queueing theory terminology, our model corresponds to a G/G/1/1 system. An example of such a G/G/1/1 system appears in the multicast problem of [24], where a new update is generated when a percentage of the destinations has received the current update, and service time to each destination is a shifted exponential random variable. Although an exact expression for their model is derived in [24], in general, calculating an exact age expression for non-exponential interarrival times is difficult. For example, [25] considers a two-stage multicast extension of [24], where only an upper bound is derived for the age of the second stage nodes. In this paper, we derive exact age expressions when the distribution of service times is arbitrary but known.

We consider two service disciplines. The first one is called G/G/1/1 *with blocking*, where a new arrival is blocked if the server is busy. This server model is memory efficient and a good choice when the server does not have a buffer. It is also used in [8] for an M/M/1/1 system and in [10] for an M/G/1/1 system. Here, we do not restrict ourselves to exponential interarrival times or exponential service times. Our first contribution is to derive an exact expression for average age in a G/G/1/1 system. This expression can be calculated using probability density functions of interarrival and service times. Next, we calculate average age expressions for M/G/1/1 and G/M/1/1 systems. Age for M/G/1/1 systems is previously derived in [10]; in this paper, we provide an alternative proof using our approach. On the other hand, average age for G/M/1/1 systems is a new contribution.

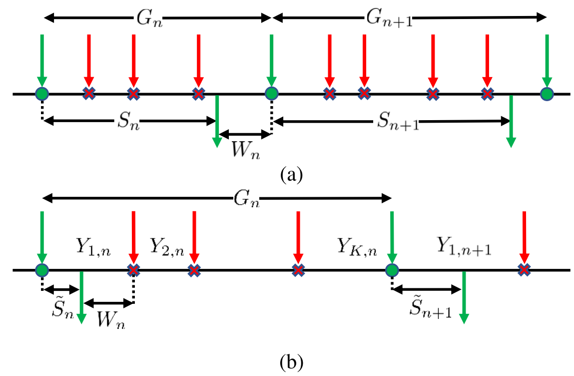


Fig. 1. Arrival and departure structure for a server. The arrows above and below the horizontal timeline corresponds to arrivals to and departures from the server. Circles are successful arrivals, while crosses are discarded arrivals: (a) blocking discipline, (b) preemption in service discipline.

For G/M/1/1 systems, we observe that deterministic interarrival times minimize the average age for a given exponential service time.

Our second service discipline model is called G/G/1/1 *with preemption in service*, where a new arrival preempts any ongoing service. This model is used in [5] and [2] for an M/M/1/1 system and in [10] for an M/G/1/1 system. Here, in this model as well, we do not restrict ourselves to exponential interarrival times or exponential service times. Our first contribution in this service discipline is to derive an exact expression for average age in a G/G/1/1 system. Unlike the case with blocking discipline, the average age in this model does not include any calculation of infinite sums. The age expression can be calculated relatively easily using probability density functions of interarrival and service times. Next, we calculate average age expressions for M/G/1/1 and G/M/1/1 systems. Age for M/G/1/1 systems is previously derived in [10]; in this paper, we provide an alternative proof using our approach. On the other hand, age for G/M/1/1 systems is a new contribution. Moreover, we prove that in a G/M/1/1 system deterministic interarrivals are optimum. We observe for non-exponential service times that the optimal distribution of interarrival times depend on the relative values of the mean interarrival time and the mean service time. Finally, we propose a simple upper bound to the average age, which does not have any restrictions on the distributions of interarrival and service times.

II. System Model

We consider a communication scenario where the data arrive at the source according to an arrival process with independent and identically distributed (i.i.d.) interarrival times Y_n. The source transmits the data through a single bufferless server. Time duration of service is modeled as a random process with i.i.d. service times S_n. Interarrival times, Y_n, and service times, S_n, are independent. We specify general probability distributions for the interarrival times and service times. Fig. 1(a) and Fig. 1(b) show realizations of the arrival/departure processes with blocking and preemption in service disciplines, respectively.

A. Blocking Discipline

In this model, if an update arrives while the server is busy, it is blocked (see cross marked arrows in Fig. 1(a)). If an update arrives while the server is idle, service is initiated immediately (see circle marked arrows in Fig. 1(a)). We refer to those updates that initiate a service as the successful updates. After a service, S_n, is completed, a successful update departs the server (see the arrow below the timeline in Fig. 1(a)). Service idle time, W_n, is the time between a departure of a successful update and the arrival of the next update. Interarrival times between consecutive successful updates, $G_n = S_n + W_n$, are called effective interarrival times. It is important to note that, the effective interarrival time, G_n, can be written as a random sum of random numbers, $G_n = \sum_{k=1}^{K} Y_{k,n}$. Note that n is the index successful updates, k is the index of arrivals between two consecutive successful updates and K is an integer random variable that describes the total number of arrivals before the next successful arrival. Probability mass function of K can be written as

$$\Pr(K = k) = \Pr\left(\sum_{j=1}^{k-1} Y_{j,n} \leq S_n < \sum_{j=1}^{k} Y_{j,n}\right). \quad (2)$$

B. Preemption in Service Discipline

In this model, if an update arrives while the server is idle, the service is initiated immediately. If an update arrives while the server is busy, the packet being served is terminated and the new packet is pushed to the server. In Fig. 1(b), each arrival (arrows above the timeline) starts a service. An arrival that can finish service is called a successful arrival (see circle marked arrows in Fig. 1(b)). Since the service time of a successful arrival needs to be smaller than the interarrival time, the time that a successful arrival stays in service is $\tilde{S}_n = S_n | S_n < Y_{1,n}$.

Similar to the blocking model, interarrival times between successful updates are called effective interarrival times, G_n, which can be written as a random sum of random numbers, $G_n = \sum_{k=1}^{K} Y_{k,n}$. Unlike the blocking model, here the waiting time depends only on the current interarrival time, $W_n = Y_{1,n} - \tilde{S}_n$. Although K in the blocking discipline does not follow a specific distribution, K in the preemption in service discipline is a geometric random variable. An effective interarrival time is the sum of a successful arrival with probability $p = \Pr(Y_{1,n} > S_n)$, and $K - 1$ unsuccessful arrivals, all with the same probability $1 - p$.

III. G/G/1/1 With Blocking

For a stationary ergodic status update system, the average age can be calculated using a geometric approach [2, Theorem 3]. For G/G/1/1 with blocking discipline, average age can be written as the difference of the areas of two triangles, divided by the expected value of the effective interarrival time. From Fig. 2, we have

$$\Delta^b_{\text{G/G}} = \frac{E[(G_n + S_{n+1})^2] - E[(S_{n+1})^2]}{2E[G_n]} \quad (3)$$

$$= \frac{E[G^2]}{2E[G]} + E[S], \quad (4)$$

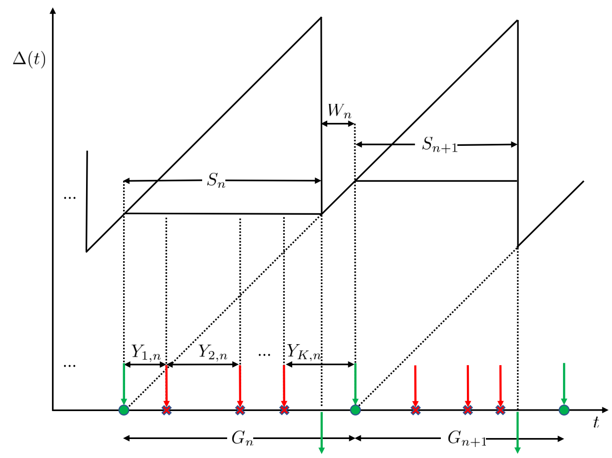


Fig. 2. Age curves for G/G/1/1 with blocking model.

where superscript b denotes *blocking*, S_{n+1} is independent of G_n, and time indices are dropped. Next, we make a general remark about (4). The effective interarrival time, G, is the time of the renewal cycle of a renewal process. We see from Fig. 2 that each time a service starts, the effective interarrival process is renewed. In [26, page 136], *average age of a renewal process* is defined, and is calculated as $\frac{E[G^2]}{2E[G]}$. Therefore, the average age of an information update in (4) is equal to the sum of the average age of the effective interarrival process and the average service time.

For most general interarrival and service time models, it is not easy to calculate the first and second moments of effective interarrival times G needed in (4). In this section, we first derive an exact expression for (4) that depends only on the general distributions of interarrival times, Y, and service times, S. Although it is exact, the average age expression for the case of general interarrival and service time distributions require further calculations on the distribution functions of the interarrival and service times. Next, we derive simpler average age expressions for several special cases, i.e., for general interarrival and exponential service times, and exponential interarrival and general service times.

A. Age for General Interarrival and Service Times

In this section, we derive an exact age expression for the case of general interarrival and service time distributions under blocking discipline. An important aspect of this result is that the average age of an information update can be written in terms of the average age of the update arrival process, $\frac{E[Y^2]}{2E[Y]}$, instead of the average age of the effective interarrival process, $\frac{E[G^2]}{2E[G]}$.

Theorem 1: Consider a G/G/1/1 system with blocking discipline, where Y_n are i.i.d. interarrival times with a general distribution and S_n are i.i.d. service times with a general distribution. The average age of an information update in this system is

$$\Delta^b_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{\sum_{k=1}^{\infty} E[A_k \bar{F}_S(A_k)]}{1 + \sum_{k=1}^{\infty} E[\bar{F}_S(A_k)]} + E[S] \quad (5)$$

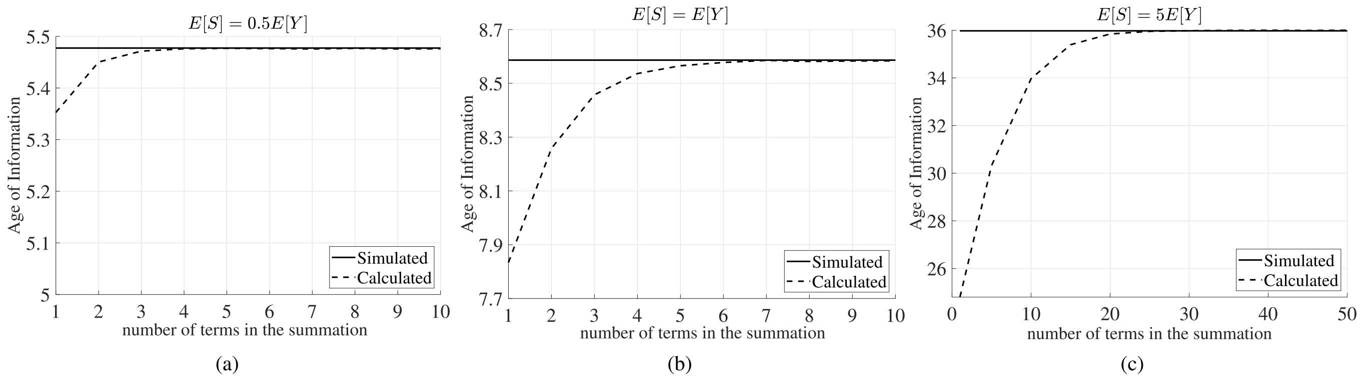


Fig. 3. AoI for G/G/1/1 with blocking discipline, where the x-axis is the number of summations that are calculated in (5) for different expected service times: (a) $\alpha_S = \alpha_Y = 2, \mu = 2, \lambda = 1$, (b) $\alpha_S = \alpha_Y = 2, \mu = \lambda = 2$, (c) $\alpha_S = \alpha_Y = 2, \mu = 0.4, \lambda = 2$.

where $A_k = \sum_{j=1}^{k} Y_j$, and $\bar{F}_S(\cdot)$ is the complementary cdf of S.

The proof of Theorem 1 is given in Section V-A. Note that Theorem 1 gives the average age of a G/G/1/1 system as a summation of three terms. The first and the third terms depend only on the interarrival and service time distributions, respectively, and the second term depends on both distributions.

The exact average age expression in (5) requires calculations of infinite sums. In Fig. 3, we consider gamma distributed[1] interarrival and service times and observe that the number of terms needed in the summation for calculation to converge depends on the ratio of the expected service time, $E[S]$, to the expected interarrival time, $E[Y]$. When $E[S] = E[Y]$, we observe from Fig. 3(b) that the error in calculation is less than 0.5% after 5 terms in the summation. In Fig. 4, we consider gamma distributed interarrival times where α_Y is the shape parameter and λ is the rate parameter, and gamma distributed service times where α_S is the shape parameter and μ is the rate parameter. Using simulations to calculate the average age, we observe the effects of the rate and shape parameters of gamma distribution on the average age. We observe that the age decreases monotonically with the rate parameter and increases monotonically with the shape parameter of either distribution.

B. Age for General Interarrival and Exponential Service Times

For exponential service times, we have a closed form expression for the average age. First, we show that for exponentially distributed service times with rate parameter μ, K is a geometric random variable with $p = 1 - E[e^{-\mu Y}]$.

Lemma 1: Consider a G/M/1/1 system with blocking discipline, where Y_n are i.i.d. interarrival times with a general distribution, S_n are i.i.d. service times with an exponential distribution with rate parameter μ, and K_n are i.i.d. discrete

[1]We choose gamma distribution in this paper in order to simulate general distributions. The gamma distribution forms a two-parameter exponential family. When the shape parameter of a gamma distribution is larger than one, it is log-concave; when the shape parameter is smaller than one, it is log-convex. The gamma distribution includes the chi-squared, Erlang, and exponential distributions as special cases. Probability density function of a gamma distribution has a flexible shape so that it can be used to approximate many probabilistic models. Log-concave distributions have important implications for age analysis. Let us consider a random variable that corresponds to the age of a device. If the random variable has a log-concave distribution, then it has an increasing probability of failure in the next instant of time, as the device ages [27]. In other words, random variables with log-concave distributions "wear out". As it happens, many common probability distributions that appear in arrival processes in real-world systems are log-concave, including exponential, Rayleigh, Erlang, gamma with shape parameter larger than one, and uniform distributions [27].

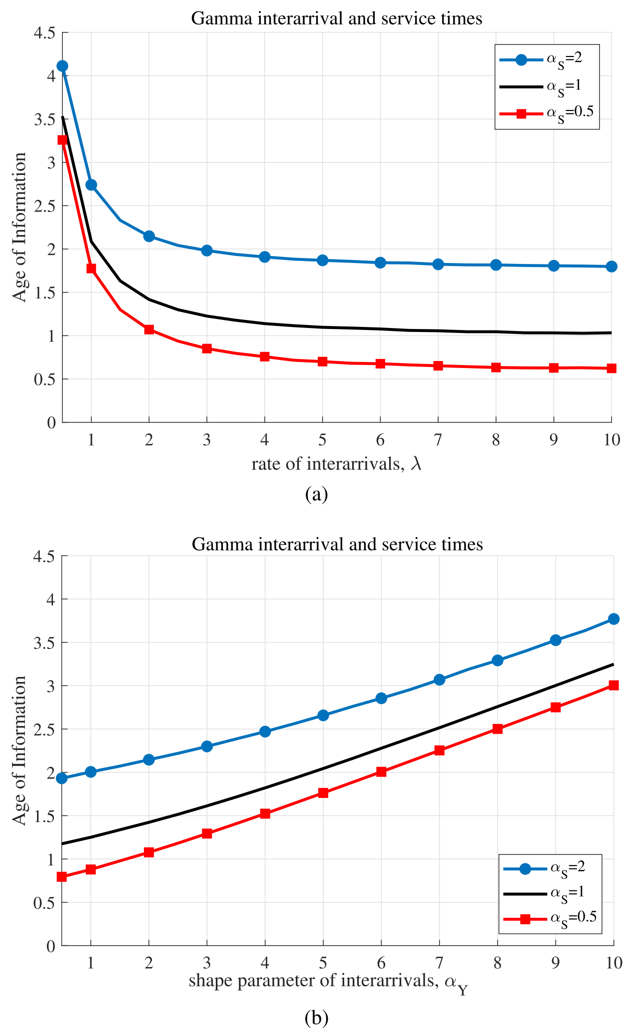


Fig. 4. AoI for G/G/1/1 with blocking discipline, where both the interarrival and service times are gamma distributed, (a) $\alpha_Y = \mu = 2$, (b) $\lambda = \mu = 2$.

random variables with

$$\Pr(K=k)=\Pr\left(\sum_{j=1}^{k-1}Y_j\leq S<\sum_{j=1}^{k}Y_j\right). \quad (6)$$

Then, K is geometric with $p=1-E[e^{-\mu Y}]$.

The proof of Lemma 1 is given in Section V-B. Using this lemma, in the next theorem, we derive the average age of a G/M/1/1 system for a given distribution for the interarrival times.

Theorem 2: Consider a G/M/1/1 system with blocking discipline, where Y_n are i.i.d. interarrival times with a general distribution and S_n are i.i.d. service times with an exponential distribution with rate parameter μ. The average age of this system is

$$\Delta^b_{\text{G/M}}=\frac{E[Y^2]}{2E[Y]}+\frac{E[Ye^{-\mu Y}]}{1-E[e^{-\mu Y}]}+\frac{1}{\mu}. \quad (7)$$

The proof of Theorem 2 is given in Section V-C. Note that similar to Theorem 1 in G/G/1/1 case, Theorem 2 gives the average age of a G/M/1/1 system as a summation of three terms. The first and the third terms depend only on the interarrival and service time distributions, respectively, and the second term depends on both distributions (note the second term has μ in it even though the expectation is with respect to Y). Unlike the G/G/1/1 case, the middle term in the average age of a G/M/1/1 system in (7) does not contain any summations. Therefore, the average age of a G/M/1/1 system can be calculated rather easily given the distribution of interarrival times. For example, when the interarrival times are also exponential with rate parameter λ, (7) reduces to

$$\Delta^b_{\text{M/M}}=\frac{1}{\lambda}+\frac{2}{\mu}-\frac{1}{\lambda+\mu} \quad (8)$$

which is the age of an M/M/1/1 queue that is found in [8].

In the next corollary, we propose an equivalent average age expression for a G/M/1/1 system with blocking discipline given in Theorem 2.

Corollary 1: The average age in Theorem 2 can be written equivalently as

$$\Delta^b_{\text{G/M}}=\frac{E[Y^2]}{2E[Y]}+2E[S]-E[S|S<Y]. \quad (9)$$

The proof of Corollary 1 is given in Section V-D. When Y is exponential, this corollary directly gives the average age of an M/M/1/1 system, since the random variable $S|S<Y$ is exponential with rate parameter $\lambda+\mu$ when Y and S are exponential with parameters λ and μ, respectively. However, we observe that calculating the average age for a general Y using Theorem 2 is easier.

In Fig. 5, we consider gamma distributed interarrival times where α_Y is the shape parameter and λ is the rate parameter, and exponential service times where μ is the rate parameter. In Fig. 5(a), we plot the average age with respect to λ when α_Y and μ are fixed, and in Fig. 5(b), we plot the average age with respect to μ when α_Y and λ are fixed. We observe that the age decreases with both rate parameters. Fig. 5(a) shows that smaller α_Y results in a lower age. In addition, for a fixed mean interarrival time ($\frac{\alpha_Y}{\lambda}$), average age is smaller

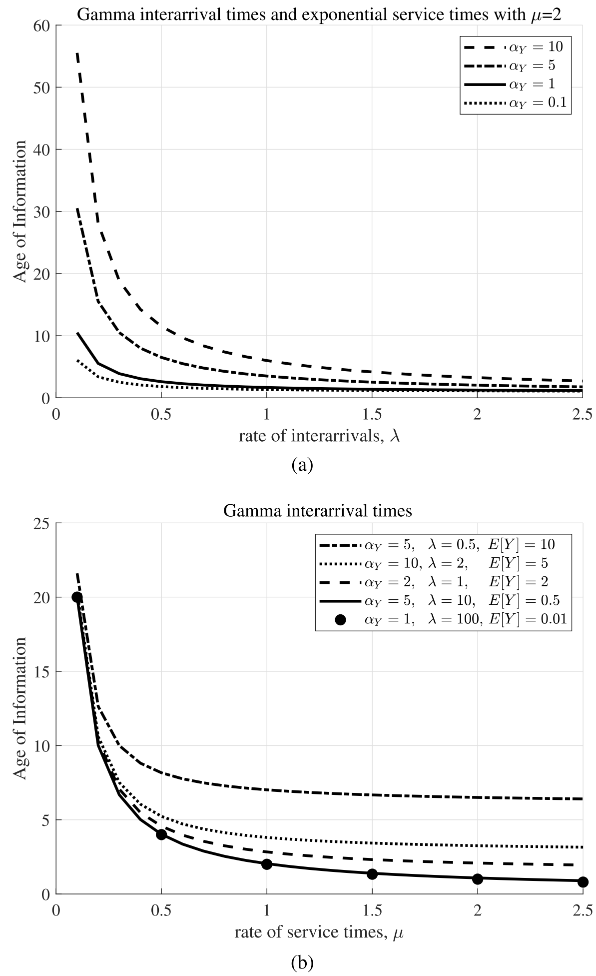


Fig. 5. AoI for G/M/1/1 with blocking discipline (a) with respect to λ when $\mu=2$ and for several α_Y, (b) with respect to μ for several (α_Y, λ) pairs.

for a smaller interarrival time variance. Similarly, for a fixed interarrival time variance ($\frac{\alpha_Y}{\lambda^2}$), average age is smaller for a smaller mean interarrival time. We see from Fig. 5(b) that the lowest age is achieved when the mean interarrival time is the smallest.

At this point, it is natural to ask what the age minimizing interarrival time distribution is for a G/M/1/1 system with a given mean interarrival time. The first term on the right hand side of (7) suggests that a distribution with the smallest second moment, which belongs to a deterministic random variable, would minimize the age. However, the effect of the middle term on the right hand side of (7) is not immediately clear. In Fig. 6, we plot (7) for several distributions. We observe that deterministic interarrivals result in the minimum age for a given mean expected interarrival time. In addition, we observe that exponential interarrivals are the worst among log-concave distributions in terms of the resulting age. Note that gamma distributions with $\alpha<1$ are not log-concave.

C. Age for Exponential Interarrival and General Service Times

In this section, we consider exponential interarrival and general service times, i.e., an M/G/1/1 system, with blocking discipline. This system is considered in [10] as well. Let us

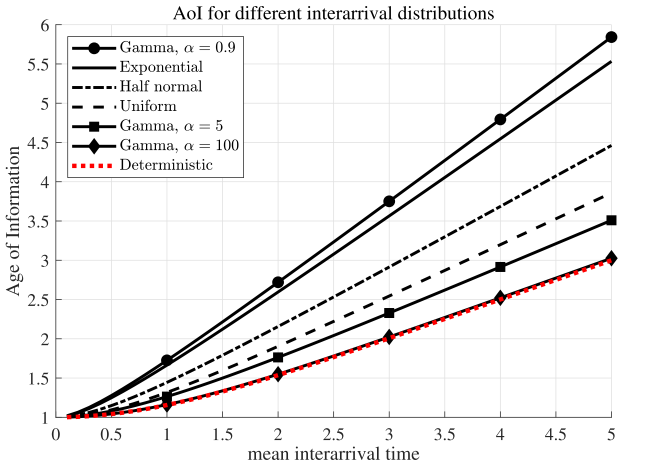


Fig. 6. AoI for G/M/1/1 with blocking discipline with respect to mean interarrival time for several interarrival distributions.

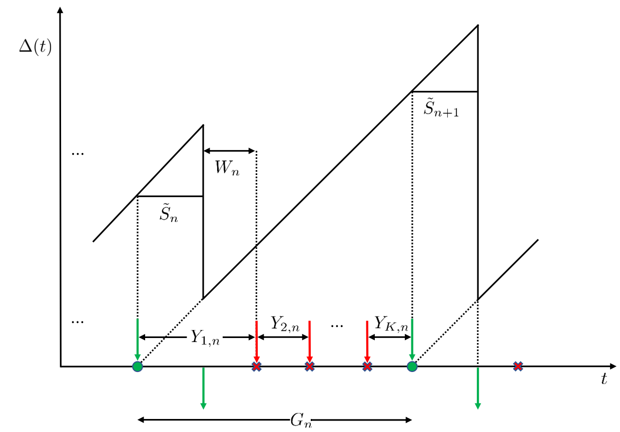


Fig. 7. Age curves for G/G/1/1 with preemption in service.

give the outline of the derivation in [10] starting from (4). Note that when interarrivals are exponential, the waiting time after a service is completed is also exponential. Therefore $G = Y + S$. Then, using the independence of Y and S, we have

$$\Delta^b_{\text{M/G}} = \frac{E[(Y+S)^2]}{2(E[Y]+E[S])} + E[S] \tag{10}$$

$$= \frac{\frac{2}{\lambda^2} + \frac{4E[S]}{\lambda} + 2E^2[S] + E[S^2]}{2(\frac{1}{\lambda} + E[S])} \tag{11}$$

$$= \frac{1}{\lambda} + \frac{\lambda E[S^2]}{2(1+\lambda E[S])} + E[S] \tag{12}$$

where $E[Y] = \frac{1}{\lambda}$ and $E[Y^2] = \frac{2}{\lambda^2}$. In Theorem 3 below, we provide an alternative proof to (12) that calculates the age expression in Theorem 1 for an exponential Y. Although this is a replication of a previous result, Theorem 3 shows how our general G/G/1/1 result in Theorem 1 can be used for specific cases.

Theorem 3: Consider an M/G/1/1 system with blocking discipline, where Y_n are i.i.d. exponential interarrival times with rate parameter λ and S_n are i.i.d. service times with a general distribution. The average age of this system is

$$\Delta^b_{\text{M/G}} = \frac{1}{\lambda} + \frac{\lambda E[S^2]}{2\,(1+\lambda E[S])} + E[S]. \tag{13}$$

The proof of Theorem 3 is given in Section V-E. We refer the reader to [10] for further analysis of M/G/1/1 systems.

IV. G/G/1/1 With Preemption in Service

Similar to the case with blocking discipline, for G/G/1/1 with preemption in service discipline as well, average age can be written as the difference of the areas of two triangles, divided by the expected length of the effective interarrival time [2, Theorem 3]. From Fig. 7, we have

$$\Delta^p_{\text{G/G}} = \frac{E[(G_n + \tilde{S}_{n+1})^2] - E[(\tilde{S}_{n+1})^2]}{2E[G_n]} \tag{14}$$

$$= \frac{E[G^2]}{2E[G]} + E[\tilde{S}] \tag{15}$$

where superscript p denotes *preemption*, $\tilde{S}_{n+1} = \{S | S < Y\}_{n+1}$ is independent of G_n, and time indices are dropped.

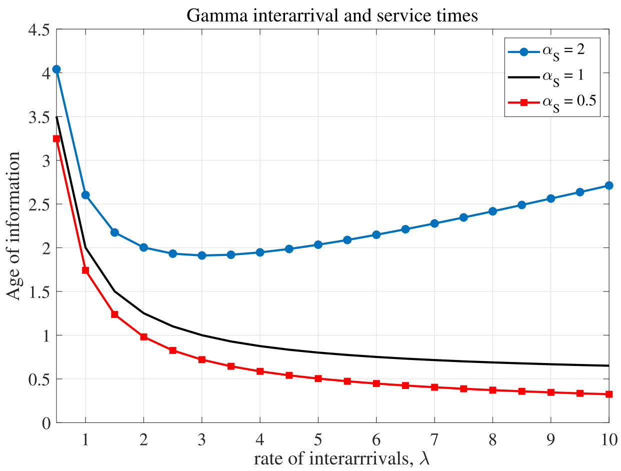
It is important to note that the random variable G in this model is not the same as the G in the blocking model. The difference can be observed from Figs. 2 and 7 by noting the change in scale for S_n. However, G represents the effective interarrival time and $\frac{E[G^2]}{2E[G]}$ represents the average age of effective interarrival process [26, page 136] in this model as well. Therefore, the average age of an information update in (15) is the sum of the average age of the effective interarrival process and the amount of time update spends in service, which is different than the mean service time of the server.

In this section, we first derive an exact closed form expression for (15) that depends only on the general distributions of interarrival times, Y, and service times, S. Unlike the age expression in the blocking model, the age expression in the preemption in service model does not require further calculations. Next, we derive simpler average age expressions for special cases, i.e., for general interarrival and exponential service times, and exponential interarrival and general service times. Finally, we derive an upper bound for the case of general interarrival and service time distributions.

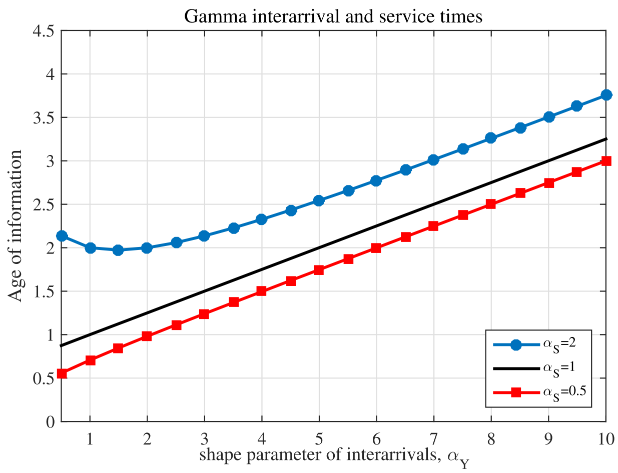
A. Age for General Interarrival and Service Times

In this section, we derive an exact age expression for the case of general interarrival and service time distributions under the preemption in service discipline. An important aspect of this result is that, unlike the case with blocking discipline, the average age of an information update in this model does not involve any infinite sums. In addition, similar to the case with blocking discipline, the average age of an information update can be written in terms of the average age of the update arrival process, $\frac{E[Y^2]}{2E[Y]}$, instead of the average age of the effective interarrival process, $\frac{E[G^2]}{2E[G]}$.

Theorem 4: Consider a G/G/1/1 system with preemption in service discipline, where Y_n are i.i.d. interarrival times with a general distribution and S_n are i.i.d. service times with a general distribution. The average age of an information update

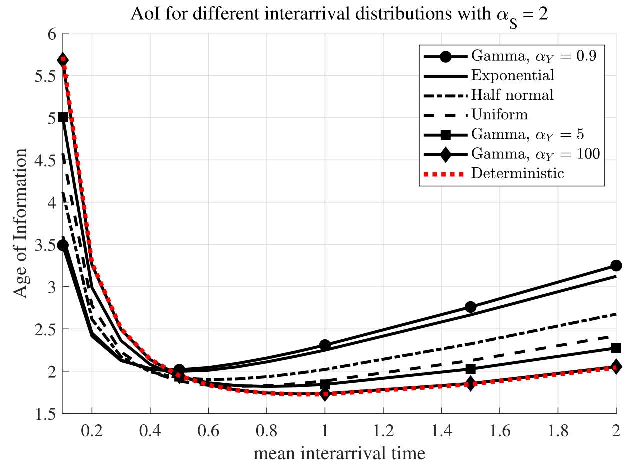


(a)

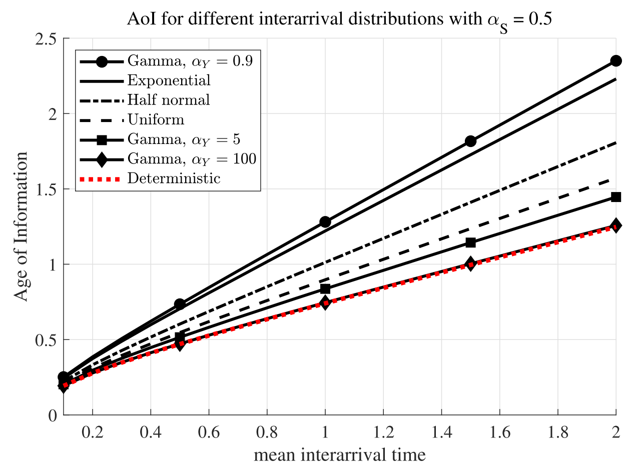


(b)

Fig. 8. AoI for G/G/1/1 with preemption in service discipline, where both the interarrival and service times are gamma distributed, (a) $\alpha_Y = \mu = 2$, (b) $\lambda = \mu = 2$.



(a)



(b)

Fig. 9. AoI for G/G/1/1 with preemption in service discipline with respect to mean interarrival time for several interarrival distributions.

in this system is

$$\Delta^p_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{E[Y\bar{F}_S(Y)]}{1-E[\bar{F}_S(Y)]} + E[\tilde{S}] \tag{16}$$

where $\bar{F}_S(\cdot)$ is the complementary cdf of S, and $\tilde{S} = S|S < Y$.

The proof of Theorem 4 is given in Section V-F.[2] Similar to Theorem 1 in the case of blocking discipline, Theorem 4 in the case of preemption in service discipline gives the average age of a G/G/1/1 system as a summation of three terms. However, unlike Theorem 1 which includes infinite sums, Theorem 4 is much easier to calculate given the distributions of interarrival and service times.

In Fig. 8, we consider gamma distributed interarrival times where α_Y is the shape parameter and λ is the rate parameter, and gamma distributed service times where α_S is the shape parameter and μ is the rate parameter. Using simulations to calculate the average age, we observe the effects of the rate and shape parameters of gamma distribution on the average age.

We have seen from Fig. 4 through Fig. 6 that the age in the blocking discipline is monotone with respect to the parameters of the gamma distribution. We observe from Fig. 8 that this is not the case in the preemption in service discipline. When λ is very large, in other words when the interarrivals are too frequent, preemption starts to overload the system when service time distribution is log-concave, i.e., $\alpha_S > 1$. Time duration between two successive successful interarrivals gets larger, and hence age increases. This observation for G/G/1/1 systems with preemption in service differs significantly from M/M/1/1 systems with preemption in service, where age is monotonically decreasing in λ [2] (see also the unmarked curves in Fig. 8(a)). In addition, the minimum age for G/G/1/1 systems over the rate parameter is smaller in the blocking scenario than it is in the preemption in service scenario. However, we know from [2] and [8] that the opposite is true for M/M/1/1 systems (see also the unmarked curves in Figs. 4(a) and 8(a)). These observations reassure our initial motivation to consider the AoI for G/G/1/1 systems, as they can behave significantly different than M/M/1/1 systems.

In Fig. 9(a), we plot the average age for different interarrival time distributions when the service time is gamma with shape parameter $\alpha_S = 2$, and rate parameter $\mu = 2$. Unlike the case

[2]A similar and independent result is also given in [14] for the average age of LCFS G/G/1 queues with preemption in service discipline. Our expression in Theorem 4 and the expression in [14] can be converted to each other using some probabilistic identities.

in the blocking discipline, deterministic interarrival times do not result in the minimum age for all mean interarrival values. We observe from Fig. 9(a) that there is a threshold, above which deterministic interarrivals are optimum and below which exponential interarrivals are optimum when the optimization is over log-concave distributions. On the other hand, in Fig. 9(b), we plot the same interarrival distributions for a gamma service time with shape parameter $\alpha_S = 0.5$, and rate parameter $\mu = 2$. Here, irrespective of the mean interarrival time, deterministic interarrivals are always optimum.

B. Age for General Interarrival and Exponential Service Times

In this section, we consider general interarrival and exponential service times in the preemption in service discipline. We derive the exact age expression that can be written as a summation of two terms, the first of which depends only on the first and second moments of interarrival times and the second of which depends only on the service rate. In other words, interarrival and service times are decoupled. Average age takes a simple representation as the sum of the age of the arrival process and the mean service time. In addition, it is interesting to see that the time an update spends in service, $\tilde{S} = S|S < Y$, disappears from the age expression.

Theorem 5: Consider a G/M/1/1 system with preemption in service, where Y_n are i.i.d. interarrival times with a general distribution and S_n are i.i.d. exponential service times with rate parameter μ. The average age of this system is

$$\Delta^p_{\text{G/M}} = \frac{E[Y^2]}{2E[Y]} + \frac{1}{\mu}. \tag{17}$$

The proof of Theorem 5 is given in Section V-G. The age expression in Theorem 5 is so simple that it only depends on the first moment of the service time, and first and second moments of the interarrival time. When the interarrivals are exponential as well, (17) reduces to

$$\Delta^p_{\text{M/M}} = \frac{1}{\lambda} + \frac{1}{\mu} \tag{18}$$

which is derived in [2].

We remark that the average age in (17) for preemption in service discipline is smaller than that in (7) for blocking discipline for all interarrival distributions. This result is also reported in [11] that shows that over all service disciplines and for exponential service times, preemption in service discipline is optimal. We remind that this result does not generalize when service times are not exponential (see the discussion in Section IV-A).

Instead of optimizing age over service disciplines as in [11], it is also possible to optimize age over the distributions of interarrival and service times. In Corollary 2, we show that deterministic interarrival times minimizes average age while exponential interarrival times maximizes the minimum age.

Corollary 2: Consider a G/M/1/1 system with preemption in service discipline, where Y_n are i.i.d. interarrival times with a general distribution, and S_n are i.i.d. exponential service times. Deterministic interarrival times are age minimum over the space of distributions with fixed mean. In addition, exponential interarrival times result in the worst age when interarrival distribution is log-concave.

Proof: For a given $E[Y]$, (17) is minimized when the second moment, or equivalently, the variance of Y is minimized. Since deterministic variables have zero variance, deterministic interarrival times minimize the average age.

When Y has a log-concave distribution, we have the following relation between the first and second moments of Y from [27, Proposition 6.A.6]

$$\frac{E[Y^2]}{2E^2[Y]} \leq 1 \tag{19}$$

where the equality is achieved with an exponential distribution. Therefore, we conclude that exponential interarrivals result in the largest possible $\frac{E[Y^2]}{2E[Y]}$ when Y has a log-concave distribution. ■

Exponential distribution results in the worst age among all log-concave distributions due to its memoryless property. A consequence of memoryless property is that the mean residual life [27] of exponential distribution is the same as its mean. For the case of preemption in service discipline, mean residual life is related to how much the server waits in idle mode after a service is completed. The longer this waiting time is, the larger the average age will be. For any other log-concave distribution, mean residual life is smaller than the mean of the distribution. Therefore, when the mean of the distribution is given and fixed, exponential distribution has the largest mean residual life. Another consequence of memoryless property is that the coefficient of variation (CV) of exponential distribution is equal to one. CV is the ratio of the standard deviation to the mean, or equivalently, the square root of the left hand side of (19). It is a measure of dispersion of a probability distribution. CV of the exponential distribution is the largest (equals to 1) among all log-concave distributions, and the smallest for deterministic distributions (equals to 0). From this, we can conclude that the minimum average age for preemption in service discipline is higher for distributions with higher dispersion.

In Fig. 10, we consider gamma distributed interarrival times where α_Y is the shape parameter and λ is the rate parameter, and exponential service times where μ is the rate parameter. In Fig. 10(a), we plot the average age with respect to λ when α_Y and μ are fixed, and in Fig. 10(b), we plot the average age with respect to μ when α_Y and λ are fixed. We observe that the age decreases with both rate parameters. Fig. 10(a) shows that smaller α_Y, smaller mean interarrival time and smaller interarrival time variance result in a smaller average age similar to the case in blocking discipline (see Fig. 5). In addition, we observe from Fig. 10(b) that the lowest age is achieved when the mean interarrival time is the smallest.

C. Age for Exponential Interarrival and General Service Times

In this section, we consider exponential interarrival and general service times, i.e., an M/G/1/1 system, with preemption in service discipline. This system is considered in [10] as well. In Theorem 6, we provide an alternative proof to [10]

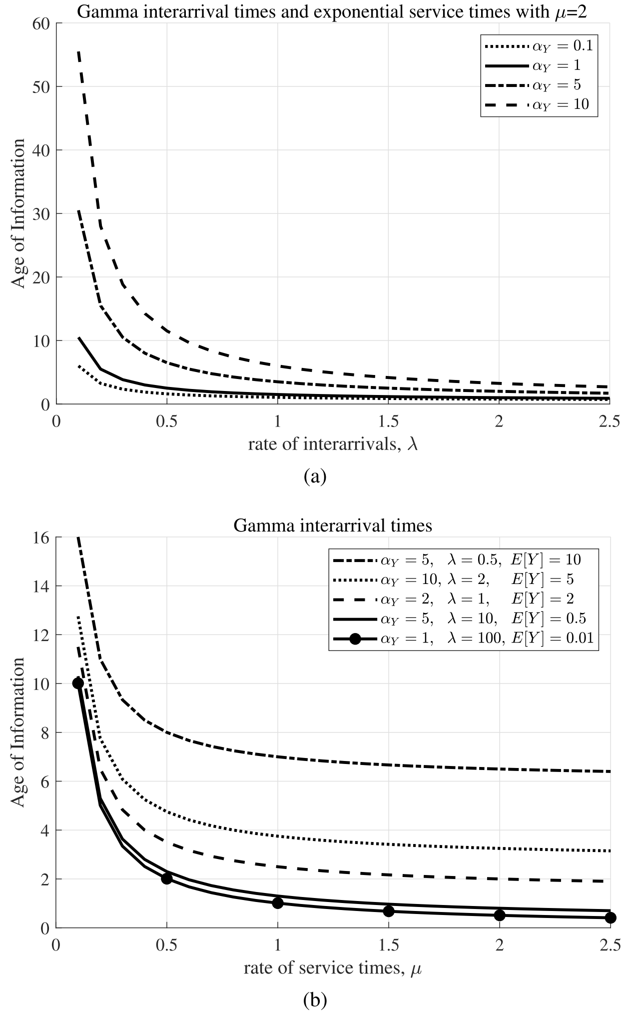


Fig. 10. AoI for G/M/1/1 with preemption in service discipline (a) with respect to λ when $\mu = 2$ and for several α, (b) with respect to μ for several (α, λ) pairs.

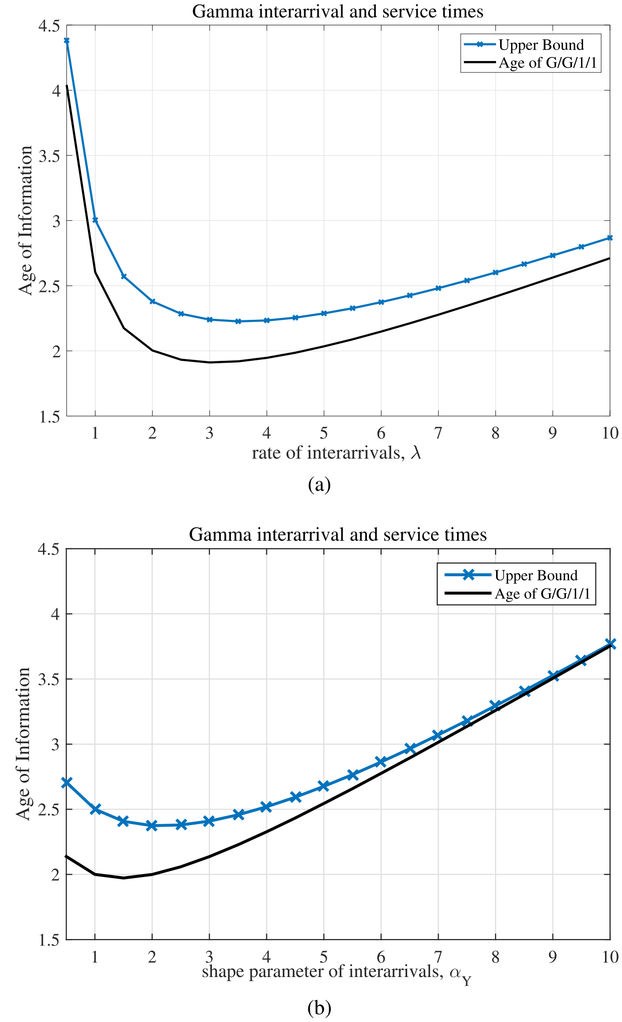


Fig. 11. AoI for G/G/1/1 with preemption in service discipline, where both the interarrival and service times are gamma distributed, (a) $\alpha_Y = \mu = 2$, (b) $\lambda = \mu = 2$.

that determines the age expression in Theorem 4 for an exponential Y. Although this is a replication of a previous result, it provides another affirmation that our approach is applicable for many different cases.

Theorem 6: Consider an M/G/1/1 system with preemption in service, where Y_n are i.i.d. exponential interarrival times with rate parameter λ and S_n are i.i.d. service times with a general distribution. The average age of this system is

$$\Delta^p_{\text{M/G}} = \frac{1}{\lambda E[e^{-\lambda S}]}. \tag{20}$$

The proof of Theorem 6 is given in Section V-H. We refer the reader to [10] for further analysis of M/G/1/1 systems.

D. Upper Bound for General Interarrival and Service Times

Our results in the previous sections provide exact age expressions for single server and single packet in the system queues with preemption in service discipline under different assumptions on the distributions of interarrival and service times. Theorem 4 gives the most general result for a G/G/1/1 system. In Corollary 3 below, we provide a simple upper bound to (16) in Theorem 4.

Corollary 3: Consider a G/G/1/1 system with preemption in service, where Y_n are i.i.d interarrival times with a general distribution and S_n are i.i.d. service times with a general distribution. The average age of this system is upper bounded by

$$\Delta^p_{\text{G/G}} \leq \frac{E[Y^2]}{2E[Y]} + \frac{E[Y]E[\bar{F}_S(Y)]}{1 - E[\bar{F}_S(Y)]} + E[\tilde{S}] \tag{21}$$

where $\bar{F}_S(\cdot)$ is the complementary cdf of S, and $\tilde{S} = S|S < Y$.

The proof of Corollary 3 follows by noting that $E[Y|Y < S] \leq E[Y]$ and $\bar{F}_S(Y) = E[1_{\{S>Y\}}|Y]$. The upper bound in (21) is tight when K is independent of Y_k. An example of this is the multicast model in [24], where the random sum parameter K is independent of Y_k.

In order to examine the tightness of the bound in Corollary 3 for a general case, in Fig. 11, we simulate the same G/G/1/1 system as in the case with blocking discipline, calculate its age using Theorem 4 (unmarked lines in Fig. 11) and compare it to the upper bound in Corollary 3 (cross marked lines in Fig. 11). We observe that the difference between the exact age and the upper bound is bounded and small. We also observe that the difference between the exact age and the upper bound depends on the interarrival and service time distributions.

V. Conclusion

Average age of status update systems has become an important metric and a design tool for data communications. However some previous results for status update systems with exponential interarrival or service times do not generalize to arbitrarily distributed interarrival and service times. For example, we show in this paper that preemption in service discipline is not always optimal when the service time distribution is not exponential. Considering real world wireless network applications that have non-exponential characteristics, in this paper, we considered average age of G/G/1/1 queues. For both the blocking and preemption in service disciplines, we derived the average age of G/G/1/1 queues as a function of interarrival and service time distributions. We also obtained the average age of G/M/1/1 queues as a special case. Finally, when AoI is optimized over the distribution of the interarrival times, we prove for the preemption in service discipline and numerically observe for the blocking discipline that deterministic interarrival times result in the minimum age and exponential interarrival times result in the maximum age.

Appendix

A. Proof of Theorem 1

Remember from Section II-A that effective interarrival times, G_n, can be written as random sums of random numbers; see Fig. 2. Although K is not independent of all Y_j, it is possible to calculate the expected value of G using Wald's equation [26, Theorem 3.3.2], which is stated in Lemma 2 below.

Lemma 2 (Wald's Eqn. [26, Theorem 3.3.2]): If $Y_1, Y_2, \ldots$ are i.i.d. random variables having finite expectations, and K is a stopping time for $Y_1, Y_2, \ldots$ such that $E[K] < \infty$, then

$$E\left[\sum_{k=1}^{K} Y_k\right] = E[K]E[Y]. \tag{22}$$

Using Wald's equation, we have $E[G] = E[K]E[Y]$. Next, we derive an expression for the second moment of the effective interarrival times, $E[G^2]$. Let us first define the indicator function,

$$I_k = \begin{cases} 1, & \text{if } k \le K \\ 0, & \text{if } k > K. \end{cases} \tag{23}$$

Now, we have

$$\begin{aligned} E[G^2] &= E\left[\left(\sum_{k=1}^{K} Y_k\right)^2\right] &(24)\\ &= E\left[\left(\sum_{k=1}^{\infty} Y_k I_k\right)^2\right] &(25)\\ &= \sum_{k=1}^{\infty} E[Y_k^2 I_k] + 2\sum_{k=2}^{\infty}\sum_{j=1}^{k-1} E[Y_k I_k Y_j I_j]. &(26)\end{aligned}$$

Note that, $I_k = 1$ if and only if we have not stopped after successively observing $Y_1, \ldots, Y_{k-1}$. Therefore, I_k is determined by $Y_1, \ldots, Y_{k-1}$, and is thus independent of Y_k. We have $E[Y_k^2 \; I_k] = E[Y_k^2]E[I_k]$, and $E[Y_k I_k Y_j I_j] = E[Y_k]E[I_k Y_j I_j]$, for $j < k$. Now, (26) becomes

$$E[G^2] = E[Y^2]\sum_{k=1}^{\infty} E[I_k] + 2E[Y]\sum_{k=2}^{\infty}\sum_{j=1}^{k-1} E[I_k Y_j I_j]. \tag{27}$$

First, let us calculate

$$\begin{aligned} \sum_{k=1}^{\infty} E[I_k] &= \sum_{k=1}^{\infty} \Pr(K \ge k) &(28)\\ &= \sum_{k=1}^{\infty}\sum_{j=k}^{\infty} \Pr(K = j) &(29)\\ &= \sum_{j=1}^{\infty}\sum_{k=1}^{j} \Pr(K = j) &(30)\\ &= \sum_{j=1}^{\infty} j\Pr(K = j) &(31)\\ &= E[K]. &(32)\end{aligned}$$

Next, let us calculate

$$\begin{aligned} \sum_{j=1}^{k-1} E[I_k Y_j I_j] &= \sum_{j=1}^{k-1} E[Y_j I_k] &(33)\\ &= E\left[\left(\sum_{j=1}^{k-1} Y_j\right) I_k\right] &(34)\\ &= E\left[A_{k-1}1_{\{A_{k-1}<S\}}\right] &(35)\\ &= E\left[E\left[A_{k-1}1_{\{A_{k-1}<S\}}|A_{k-1}\right]\right] &(36)\\ &= E\left[A_{k-1}E\left[1_{\{A_{k-1}<S\}}|A_{k-1}\right]\right] &(37)\\ &= E\left[A_{k-1}\bar{F}_S(A_{k-1})\right] &(38)\end{aligned}$$

where $A_{k-1} = \sum_{j=1}^{k-1} Y_j$, and we used the fact that $I_k = 1$ implies that $I_j = 1$ for every $j < k$. Now, (27) becomes

$$\begin{aligned} E\left[G^2\right] &= E\left[Y^2\right]E[K] + 2E[Y]\sum_{k=2}^{\infty} E[A_{k-1}\bar{F}_S(A_{k-1})] &(39)\\ &= E\left[Y^2\right]E[K] + 2E[Y]\sum_{k=1}^{\infty} E[A_k\bar{F}_S(A_k)]. &(40)\end{aligned}$$

Inserting (40) and $E[G] = E[K]E[Y]$ into (4), we have

$$\Delta^b_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{\sum_{k=1}^{\infty} E[A_k\bar{F}_S(A_k)]}{E[K]} + E[S]. \tag{41}$$

Now, let us write $E[K]$ in terms of S and Y as

$$\begin{aligned} E[K] &= \sum_{k=1}^{\infty} \Pr(S > A_{k-1}) &(42)\\ &= \sum_{k=0}^{\infty} \Pr(S > A_k) &(43)\\ &= 1 + \sum_{k=1}^{\infty} E[\bar{F}_S(A_k)]. &(44)\end{aligned}$$

Finally, inserting (44) into (41), we have (5).

B. Proof of Lemma 1

Let us start with $\Pr(K=k)$ in (6) as,

$$\Pr(K=k) = \Pr\left(\sum_{j=1}^{k-1} Y_j \leq S < \sum_{j=1}^{k} Y_j\right) \tag{45}$$

$$= E\left[\bar{F}_S\left(\sum_{j=1}^{k-1} Y_j\right) - \bar{F}_S\left(\sum_{j=1}^{k} Y_j\right)\right] \tag{46}$$

$$= E\left[e^{-\mu\left(\sum_{j=1}^{k-1} Y_j\right)} - e^{-\mu\left(\sum_{j=1}^{k} Y_j\right)}\right] \tag{47}$$

$$= \left(E\left[e^{-\mu Y}\right]\right)^{k-1} - \left(E\left[e^{-\mu Y}\right]\right)^{k} \tag{48}$$

$$= \left(E\left[e^{-\mu Y}\right]\right)^{k-1}\left(1 - E\left[e^{-\mu Y}\right]\right), \tag{49}$$

which shows that K is geometric with $p = 1 - E\left[e^{-\mu Y}\right]$.

C. Proof of Theorem 2

Let us start with calculating the middle term on the right hand side of (5) for an exponential S. First, note that the denominator is $E[K]$ (see (41) and (44)). Due to Lemma 1, K is geometric with $p = 1 - E\left[e^{-\mu Y}\right]$ when S is exponential with rate μ. Therefore, we have $E[K] = \frac{1}{1-E[e^{-\mu Y}]}$. Next, let us consider the numerator. We have

$$\sum_{k=1}^{\infty} E[A_k \bar{F}_S(A_k)] = \sum_{k=1}^{\infty} E\left[\left(\sum_{j=1}^{k} Y_j\right) e^{-\mu\left(\sum_{j=1}^{k} Y_j\right)}\right] \tag{50}$$

$$= \sum_{k=1}^{\infty} E\left[\left(\sum_{j=1}^{k} Y_j\right) \prod_{j=1}^{k} e^{-\mu Y_j}\right] \tag{51}$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{k} E\left[Y_j e^{-\mu Y_j} \prod_{j'\neq j}^{k} e^{-\mu Y_{j'}}\right] \tag{52}$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{k} E\left[Y_j e^{-\mu Y_j}\right]\left(E[e^{-\mu Y}]\right)^{k-1} \tag{53}$$

$$= E\left[Y e^{-\mu Y}\right] \sum_{k=1}^{\infty} k\left(E[e^{-\mu Y}]\right)^{k-1} \tag{54}$$

$$= \frac{E\left[Y e^{-\mu Y}\right]}{\left(1 - E[e^{-\mu Y}]\right)^2} \tag{55}$$

where in (55), we used the fact that $E[e^{-\mu Y}] < 1$. Since $e^{-\mu Y} < 1$ for every realization of Y except a single point, $Y = 0$, which has probability zero. Finally, inserting (55) into (5) with $E[K] = \frac{1}{1-E[e^{-\mu Y}]}$, we have (7).

D. Proof of Corollary 1

First, note that

$$E[Y|Y<S]\Pr(Y<S) = E\left[Y 1_{\{Y<S\}}\right] \tag{56}$$

$$= E\left[E\left[Y 1_{\{Y<S\}}|Y\right]\right] \tag{57}$$

$$= E\left[Y E\left[1_{\{Y<S\}}|Y\right]\right] \tag{58}$$

$$= E\left[Y \bar{F}_S(Y)\right]. \tag{59}$$

Now, consider the middle term on the right hand side of (7). For an exponential S, we have $E[Y e^{-\mu Y}] = E[Y\bar{F}_S(Y)]$, which using (59), can be written as

$$E[Y e^{-\mu Y}] = E[Y|Y<S]\Pr(Y<S) \tag{60}$$

For an exponential S, we also have $1 - E[e^{-\mu Y}] = \Pr(Y > S)$. Now, the middle term on the right hand side of (7) can be written as

$$\frac{E[Y e^{-\mu Y}]}{1 - E[e^{-\mu Y}]} = \frac{E[Y|Y<S]\Pr(Y<S)}{\Pr(Y>S)} \tag{61}$$

When S is exponential, Y is nonnegative and S is independent of Y, we have

$$E[S] = E[S - Y|S > Y] \tag{62}$$

$$= E[S|S>Y] - E[Y|S>Y] \tag{63}$$

due to the memoryless property of the exponential distribution. By pulling $E[Y|S>Y]$ from (63) and inserting it into (61), we have

$$\frac{E[Y e^{-\mu Y}]}{1-E[e^{-\mu Y}]} = \frac{E[S|S>Y]\Pr(S>Y) - E[S](1-\Pr(Y>S))}{\Pr(Y>S)} \tag{64}$$

$$= \frac{E[S|S>Y]\Pr(S>Y) - E[S] + E[S]\Pr(Y>S)}{\Pr(Y>S)} \tag{65}$$

We know that

$$E[S] = E[S|S<Y]\Pr(S<Y) + E[S|S>Y]\Pr(S>Y) \tag{66}$$

Using (66), (65) becomes

$$\frac{E[Y e^{-\mu Y}]}{1 - E[e^{-\mu Y}]} = \frac{-E[S|S<Y]\Pr(S<Y) + E[S]\Pr(Y>S)}{\Pr(Y>S)} \tag{67}$$

$$= -E[S|S<Y] + E[S] \tag{68}$$

By inserting (68) into (7) and noting that $E[S] = \frac{1}{\mu}$, we have (9).

E. Proof of Theorem 3

First, note that

$$E[A_k|A_k<S]\Pr(A_k<S) = E\left[A_k 1_{\{A_k<S\}}\right] \tag{69}$$

$$= E\left[E\left[A_k 1_{\{A_k<S\}}|A_k\right]\right] \tag{70}$$

$$= E\left[A_k E\left[1_{\{A_k<S\}}|A_k\right]\right] \tag{71}$$

$$= E\left[A_k \bar{F}_S(A_k)\right]. \tag{72}$$

Now, consider the middle term on the right hand side of (5). Using (72), the numerator can be written as

$$\sum_{k=1}^{\infty} E[A_k \bar{F}_S(A_k)] = \sum_{k=1}^{\infty} E\left[A_k|A_k<S\right]\Pr(A_k<S) \tag{73}$$

$$= \sum_{k=1}^{\infty}\int_0^{\infty} y\Pr(A_k<S|A_k=y) f_{A_k}(y)dy \tag{74}$$

$$= E_S\left[\int_0^{S}\sum_{k=1}^{\infty} y f_{A_k}(y)dy\right]. \tag{75}$$

Note that A_k is a sum of k exponentials, which has an Erlang distribution. Using the density of Erlang distribution with rate parameter λ and shape parameter k, we have

$$\sum_{k=1}^{\infty} y f_{A_k}(y) = \sum_{k=1}^{\infty} y \frac{\lambda^k y^{k-1} e^{-\lambda y}}{(k-1)!} \tag{76}$$
$$= \lambda y e^{-\lambda y} \sum_{k=1}^{\infty} \frac{\lambda^{k-1} y^{k-1}}{(k-1)!} \tag{77}$$
$$= \lambda y e^{-\lambda y} \sum_{k=0}^{\infty} \frac{(\lambda y)^k}{k!} \tag{78}$$
$$= \lambda y e^{-\lambda y} e^{\lambda y} \tag{79}$$
$$= \lambda y \tag{80}$$

where we used the power series expansion of the exponential function $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$. Now, (75) becomes

$$\sum_{k=1}^{\infty} E[A_k \bar{F}_S(A_k)] = E_S\left[\int_0^S \lambda y dy\right] \tag{81}$$
$$= \frac{\lambda E_S[S^2]}{2}. \tag{82}$$

Finally, for an exponential Y, we know that $G = Y + S$. Using the fact that $E[G] = E[K]E[Y]$, we obtain $E[K] = \frac{E[Y]+E[S]}{E[Y]}$. Inserting $E[K]$ and (82) into (5), we have (13).

F. Proof of Theorem 4

Remember from Section II-B that the effective interarrival time, $G = \sum_{k=1}^{K} Y_k$ is a random sum of random numbers, where K is a geometric random variable. From Lemma 2, we have $E[G] = E[K]E[Y]$. Next, we derive an expression for the second moment of the effective interarrival times, $E[G^2]$. Let us first use the indicator function in (23) and the expansion of $E[G^2]$ in (26). Similar to the case of blocking discipline, here, I_k is independent of Y_k as well, and therefore, we have (27). Let us consider

$$\sum_{j=1}^{k-1} E[I_k Y_j I_j] = \sum_{j=1}^{k-1} E[Y_j | I_k = 1]\Pr(I_k = 1) \tag{83}$$
$$= (k-1)E[Y | Y < S]E[I_k] \tag{84}$$

where we used the fact that $I_k = 1$ implies $Y_j < S$ for $j < k$. Now, $E[G^2]$ becomes

$$E[G^2] = E[Y^2]E[K] + 2E[Y]E[Y|Y<S]\sum_{k=2}^{\infty}(k-1)E[I_k] \tag{85}$$

where $\sum_{k=1}^{\infty} E[I_k] = E[K]$ is shown in (32). Now, let us calculate

$$\sum_{k=2}^{\infty}(k-1)E[I_k] = \sum_{k=1}^{\infty}(k-1)\Pr(K \geq k) \tag{86}$$
$$= \sum_{k=1}^{\infty}\sum_{j=k}^{\infty}(k-1)\Pr(K = j) \tag{87}$$
$$= \sum_{j=1}^{\infty}\sum_{k=1}^{j}(k-1)\Pr(K = j) \tag{88}$$
$$= \sum_{j=1}^{\infty}\Pr(K = j)\sum_{k=1}^{j}(k-1) \tag{89}$$
$$= \sum_{j=1}^{\infty}\frac{(j-1)j}{2}\Pr(K = j) \tag{90}$$
$$= \frac{1}{2}E[K(K-1)]. \tag{91}$$

Thus, we have

$$E[G^2] = E[Y^2]E[K] + E[Y]E[Y|Y < S]E[K(K-1)]. \tag{92}$$

Next, using (59), the average age can be written as

$$\Delta^p_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + E[Y\bar{F}_S(Y)]\frac{E[K(K-1)]}{2E[K](1-p)} + E[\tilde{S}]. \tag{93}$$

Since K is geometric with $p = 1 - E[\bar{F}_S(Y)]$, we have

$$\frac{E[K(K-1)]}{2E[K]} = \frac{\frac{2-p}{p^2} - \frac{1}{p}}{\frac{2}{p}} = \frac{1-p}{p}. \tag{94}$$

Inserting (94) into (93), we have (16).

G. Proof of Theorem 5

Let us start by re-writing (16) using (59) and the definition of $\tilde{S}$

$$\Delta^p_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{E[Y|Y<S]\Pr(Y<S)}{\Pr(S<Y)} + E[S|S<Y]. \tag{95}$$

We know that

$$E[S] = E[S|S<Y]\Pr(S<Y) + E[S|S>Y]\Pr(S>Y). \tag{96}$$

By pulling $E[S|S < Y]$ from (96) and inserting it into (95), we have

$$\Delta^p_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{E[Y|Y<S]\Pr(Y<S)}{\Pr(S<Y)} + \frac{E[S] - E[S|S>Y]\Pr(S>Y)}{\Pr(S<Y)} \tag{97}$$
$$= \frac{E[Y^2]}{2E[Y]} + \frac{E[S] - E[S-Y|S>Y]\Pr(S>Y)}{\Pr(S<Y)}. \tag{98}$$

When S is exponential, Y is nonnegative and S is independent of Y, $E[S] = E[S - Y|S > Y]$ due to the memoryless property of the exponential distribution. Then, (98) becomes

$$\Delta^p_{\text{G/M}} = \frac{E[Y^2]}{2E[Y]} + \frac{E[S]\left(1 - \Pr(S > Y)\right)}{\Pr(S < Y)} \tag{99}$$
$$= \frac{E[Y^2]}{2E[Y]} + E[S] \tag{100}$$

which gives (17) by noting that $E[S] = \frac{1}{\mu}$.

H. Proof of Theorem 6

Let us start by re-writing (16) as

$$\Delta^p_{\text{G/G}} = \frac{E[Y^2]}{2E[Y]} + \frac{E[Y|Y<S]\Pr(Y<S)}{\Pr(S<Y)} + E[S|S<Y]. \tag{101}$$

For an exponential Y, we can calculate

$$E[Y|Y < S]\Pr(Y < S) = \int_0^{\infty} y\Pr(Y < S|Y = y)f_Y(y)dy \tag{102}$$
$$= E_S\left[\int_0^S y\lambda e^{-\lambda y}dy\right] \tag{103}$$
$$= \frac{1 - E[e^{-\lambda S}]}{\lambda} - E[Se^{-\lambda S}]. \tag{104}$$

We know that $\Pr(S < Y) = \bar{F}_Y(S) = E[e^{-\lambda S}]$. Finally, $E[S|S<Y]$ can be calculated as

$$E[S|S<Y] = \int_0^\infty s \frac{\Pr(S<Y|S=s)}{\Pr(S<Y)} f_S(s) ds \tag{105}$$

$$= \frac{E[Se^{-\lambda S}]}{E[e^{-\lambda S}]}. \tag{106}$$

Inserting (104) and (106) into (101) gives (20).

REFERENCES

[1] A. Soysal and S. Ulukus, "Age of information in G/G/1/1 systems," in *Proc. Asilomar Conf.*, Nov. 2019.

[2] 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.

[3] F. Metzger, T. Hobfeld, A. Bauer, S. Kounev, and P. E. Heegaard, "Modeling of aggregated IoT traffic and its application to an IoT cloud," *Proc. IEEE*, vol. 107, no. 4, pp. 679–694, Apr. 2019.

[4] C. Majumdar, M. Lopez-Benitez, and S. N. Merchant, "Real smart home data-assisted statistical traffic modeling for the Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4761–4776, Jun. 2020.

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

[6] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *Proc. 46th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2012, pp. 1–6.

[7] R. D. Yates and S. Kaul, "Real-time status updating: Multiple sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2666–2670.

[8] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1897–1910, Apr. 2016.

[9] E. Najm and R. Nasser, "Age of information: The gamma awakening," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2574–2578.

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

[11] A. M. Bedewy, Y. Sun, and N. B. Shroff, "Minimizing the age of information through queues," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 5215–5232, Aug. 2019.

[12] 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," *IEEE Trans. Inf. Theory*, vol. 65, no. 12, pp. 8305–8324, Dec. 2019.

[13] J. Prakash Champati, H. Al-Zubaidy, and J. Gross, "On the distribution of AoI for the GI/GI/1/1 and GI/GI/1/2* systems: Exact expressions and bounds," 2019, *arXiv:1905.04068*. [Online]. Available: http://arxiv.org/abs/1905.04068

[14] R. Talak, S. Karaman, and E. Modiano, "Can determinacy minimize age of information?" 2018, *arXiv:1810.04371*. [Online]. Available: http://arxiv.org/abs/1810.04371

[15] 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. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.

[16] B. T. Bacinoglu, E. T. Ceran, and E. Uysal-Biyikoglu, "Age of information under energy replenishment constraints," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2015, pp. 25–31.

[17] A. Arafa and S. Ulukus, "Age-minimal transmission in energy harvesting two-hop networks," in *Proc. GLOBECOM IEEE Global Commun. Conf.*, Dec. 2017, pp. 1–6.

[18] A. Arafa and S. Ulukus, "Age minimization in energy harvesting communications: Energy-controlled delays," in *Proc. 51st Asilomar Conf. Signals, Syst., Comput.*, Oct. 2017, pp. 1801–1805.

[19] X. Wu, J. Yang, and J. Wu, "Optimal status update for age of information minimization with an energy harvesting source," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 193–204, Mar. 2018.

[20] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor, "Age-minimal online policies for energy harvesting sensors with incremental battery recharges," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2018, pp. 1–10.

[21] A. Arafa, J. Yang, and S. Ulukus, "Age-minimal online policies for energy harvesting sensors with random battery recharges," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2018, pp. 1–6.

[22] A. Baknina, O. Ozel, J. Yang, S. Ulukus, and A. Yener, "Sending information through status updates," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2018, pp. 2271–2275.

[23] 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.

[24] J. Zhong, E. Soljanin, and R. D. Yates, "Status updates through multicast networks," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 463–469.

[25] B. Buyukates, A. Soysal, and S. Ulukus, "Age of information in two-hop multicast networks," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 513–517.

[26] S. M. Ross, *Stochastic Processes*. Hoboken, NJ, USA: Wiley, 1996.

[27] A. W. Marshall and I. Olkin, *Life Distributions: Structure of Nonparametric, Semiparametric and Parametric Families*. Springer, 2007.

Alkan Soysal (Member, IEEE) received the B.S. degree in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park. He is currently the Boeing Collegiate Associate Professor of electrical and computer engineering in the Calhoun Discovery Program at Virginia Tech Honors College, where he is also affiliated with the Wireless@Virginia Tech research group within The Bradley Department of Electrical and Computer Engineering, Virginia Tech. Previously, he held a visiting associate professor position with the Department of Electrical and Computer Engineering, University of Maryland. From 2008 to 2018, he was with Bahcesehir University, Istanbul, Turkey, first as an Assistant Professor and then as an Associate Professor at the Department of Electrical and Electronics Engineering. His research interests include optimization of wireless communication networks and information theory. His current research focuses on the age of information, the Internet of Things for Industry 4.0, machine learning for wireless networks, and massive MIMO.

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 the Anthony Ephremides Professor in information sciences and systems with the Department of Electrical and Computer Engineering, University of Maryland, College Park, 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 at AT&T Labs-Research. Her research interests include information theory, wireless communications, machine learning, signal processing and networks; with recent focus on private information retrieval, age of information, distributed coded computing, machine learning for wireless, energy harvesting communications, physical layer security, and wireless energy and information transfer.

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 2005 NSF CAREER Award, the 2010–2011 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 for 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 is an Area Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (since 2019) and a Senior Editor of the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING (since 2020). She was an Area Editor for the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING (2016–2020), an Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Series on "Green communications and networking" (2015–2016), an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY (2007–2010), and an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS (2003–2007). She was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (2008, 2015 and 2021), *Journal of Communications and Networks* (2012), and the IEEE TRANSACTIONS ON INFORMATION THEORY (2011).