Age-of-Information in the Presence of Error

Kun Chen
IIS, Tsinghua University
chenkun14@mails.tsinghua.edu.cn

Longbo Huang
IIS, Tsinghua University
longbohuang@tsinghua.edu.cn

***Abstract*—We consider the peak age-of-information (PAoI) in an $M/M/1$ queueing system with packet delivery error, i.e., update packets can get lost during transmissions to their destination. We focus on two types of policies, one is to adopt Last-Come-First-Served (LCFS) scheduling, and the other is to utilize retransmissions, i.e., keep transmitting the most recent packet. Both policies can effectively avoid the queueing delay of a busy channel and ensure a small PAoI. Exact PAoI expressions under both policies with different error probabilities are derived, including First-Come-First-Served (FCFS), LCFS with preemptive priority, LCFS with non-preemptive priority, Retransmission with preemptive priority, and Retransmission with non-preemptive priority. Numerical results obtained from analysis and simulation are presented to validate our results.**

I. Introduction

Many information systems work in such a mode that status updates are first collected from a time-varying environment, and then control decisions are made based on these information. Examples include sensor networks for large-scale monitoring [1], vehicular networks where vehicle position and velocity information are disseminated to assist safe and intelligent transportation [2], and wireless networks where scheduling is carried out based on channel state information [3]. A key to these systems is to *ensure timely delivery of status updates*, since out-of-date information can lead to incorrect system status estimation and result in severe performance loss.

Age-of-information (AoI), first proposed in [4], provides a measure for the "freshness" of the current status information, and is an important metric for measuring quality-of-service (QoS) of a system. Different from typical performance metrics such as delay or throughput, AoI jointly captures the latency in transmitting updates and the rate at which they are delivered.

There have been various recent works on understanding AoI. [4] analyzes AoI for queueing models including $M/M/1$, $M/D/1$ and $D/M/1$. A more complicated case with multiple update sources is analyzed in [5]. [6] studies AoI in a Last-Come-First-Served (LCFS) $M/M/1$ queueing system with or without preemption. The case when the destination may receive out-of-order packets is considered in [7]. In [8], the authors introduce a notion *peak age-of-information* (PAoI) and consider systems with packet management, i.e., the queue can choose to only keep a subset of update packets. AoI in a multi-class $M/G/1$ queueing system is studied in [9]. In [10], the authors study optimal update scheduling in a discrete-time multi-source system. The optimal update generating policy is explored in [11].

We notice that one common assumption made in most aforementioned works is that update packet delivery is always perfect, and AoI has been investigated mostly under the First-Come-First-Served (FCFS) principle. An exception is [6], which studies AoI under the LCFS principle, but also assumes perfect packet delivery. However, in practical systems, packet transmissions often contain errors and losses, e.g., due to interference or buffer overflow at intermediate routers in a multi-hop network. To study the impact of such delivery errors on AoI, in this paper, we focus on an $M/M/1$ queueing model where each packet, upon service completion, arrives at the destination with a nonzero probability. Our model captures (i) the queueing effect, which approximates the process where update packets are sent over a channel or a network and can cause congestion (This is different from [10], which also considers transmission errors), and (ii) the error component, which models the fact that update packets can get lost during the delivery process.

We first focus on the LCFS service principle and derive the exact PAoI for both the systems with preemptive priority and non-preemptive priority. Intuitively, LCFS is good for two reasons. (i) Compared to packet management schemes, e.g, [8], LCFS similarly avoids delaying new update packets with queueing by letting them go first. This results in significant reduction of AoI compared to FCFS, especially when the channel utilization is high. (ii) When there are errors in packet transmissions, packet management schemes can suffer severely due to the lack of updates to deliver, while LCFS still ensures a good delivery rate and does not affect AoI significantly.

Next we analyze the PAoI under retransmission schemes. Here we do not assume feedback, since retransmissions based on feedback may suffer from waste of time waiting for feedback, or interference between update packets and feedback information. Thus, the Retransmission policies refer to keep transmitting the most recent packet repeatedly until a new packet arrives. Compared to LCFS, retransmission policies have an advantage of always transmitting the most recent updates, at the cost of additional packet state management. We also derive the exact PAoI expressions for retransmission with or without preemption.

In this work LCFS and Retransmission policies are both studied to cover various scenarios. Although utilizing retransmissions is expected to contribute to a small AoI, it does not apply to scenarios where transmissions are not guaranteed, e.g., UDP and some wireless sensor networks. The rest of the paper is organized as follows. In Section II we introduce the model. In Sections III, IV and V we present our analysis for the FCFS, LCFS preemptive and LCFS non-preemptive cases, while in Section VI we give the results of the Retransmission preemptive and Retransmission non-preemptive cases. In Section VII we present numerical results. We conclude the paper in Section VIII. Due to space limit, some details are omitted in this paper. Readers can refer to [12] for more details.

II. System Model

We consider a system where a source transmits updates (packets) to a remote destination through a queue. The source generates packets according to a Poisson process with rate λ. The service time for each packet is exponentially distributed with service rate μ. For simplicity, we define

$$\rho = \lambda/\mu, \quad \rho_\lambda = \lambda/(\lambda+\mu), \quad \rho_\mu = \mu/(\lambda+\mu).$$

Different from previous works, we assume that upon service completion, each packet arrives at the destination independently with probability $p \in [0,1]$. Such a system is modeled by an $M/M/1$ queueing system with packet loss, as shown in Fig. 1. The packet loss model captures real-world situations where update packets can get lost during delivery to their destination, e.g., interference or buffer overflow, and has not yet been studied.

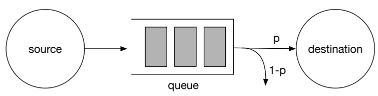


Fig. 1. The packet delivery process in a queueing system with packet loss.

We study the peak age-of-information (PAoI) in this system, which is defined as follows. Suppose each update packet has a time-stamp, marking its generation time. Denote the time-stamp of the most recently received update at time t as $\delta(t)$. Then, the status age is defined as [4]:

$$\Delta(t) \triangleq t - \delta(t),$$

and the set of peak ages is defined as:

$$\{\Delta(t_i) | \exists \epsilon > 0 \text{ s.t. } \forall t \in (t_i - \epsilon) \cup (t_i + \epsilon), \Delta(t) < \Delta(t_i)\}.$$

Then, PAoI [8] is defined to be:

$$A_P \triangleq \lim_{I\to\infty} \frac{1}{I} \sum_{i=1}^{I} \Delta_i = \mathbb{E}\{\Delta_i\},$$

where $\Delta_i = \Delta(t_i)$ is the i-th peak of $\Delta(t)$ (See Fig. 2). The last equality follows from the ergodicity of Δ_i. As shown in [9], PAoI is closely related to the average AoI, but is much more tractable.

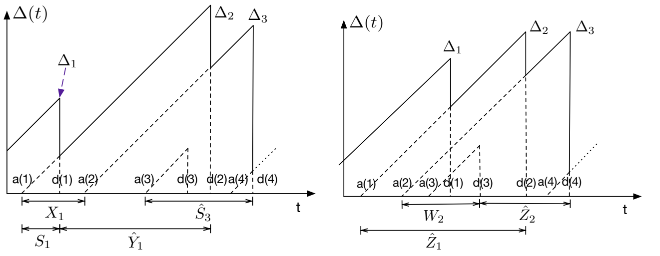
We first introduce some useful definitions. Denote N the set of all packets, according to the order in which they arrive. For a packet n, denote $a(n)$ its arrival time , $d(n)$ its departure time and $u(n)$ the time it starts to receive service. Let Φ denote the set of all successfully transmitted packets. Under the LCFS service discipline, a successfully transmitted packet may be outdated when arriving at the destination. Thus, we further define the set of *informative* packets Ψ as:

$$\Psi \triangleq \{n \in \Phi \,|\, d(n) - a(n) < \Delta(d(n))\}.$$

That is, $\Psi = \{n_1, n_2, \ldots, n_i, \ldots\}$ contains the packets which offer new information (so the system age decreases) when they reach the destination.

Regarding the evolution of the system, we define the following random variables:

$$X_n \triangleq a(n+1) - a(n), W_n \triangleq u(n) - a(n), S_n \triangleq d(n) - u(n),$$



(a) Preemptive (b) Non-preemptive

Fig. 2. Evolution of status age in the LCFS $M/M/1$ system. PAoI is divided in different ways under the preemptive and non-preemptive cases.

i.e., X_n is the inter-arrival time between n and $n+1$; W_n is the waiting time of n; S_n is the "service time" of n. Note that in the LCFS with preemptive priority case, S_n may include service time of later packets if n is preempted by other packets.

III. PAoI under FCFS

For the basic First-Come-First-Served (FCFS) case with packet loss, PAoI can be easily obtained. Since under FCFS $\Phi = \Psi$, PAoI is composed of the (expected) inter-arrival time of two successfully transmitted packets and the time a packet spends in the system. The former is $\frac{1}{p\lambda}$ and the latter is $\frac{1}{\mu-\lambda}$, resulting in a PAoI of (For rigorous derivation, see [12]):

$$A_P^{FCFS} = \frac{1}{p\lambda} + \frac{1}{\mu - \lambda}. \tag{1}$$

However, the FCFS policy, as discussed above, can suffer from traffic congestion, under which each packet will take a long time to get through the queue and the PAoI can be poor. Thus, in this work, we focus on the Last-Come-First-Served (LCFS) as well as Retransmission policies and consider the following two scheduling schemes.

1) Preemptive priority: If a new packet arrives while the server is busy, it preempts the current packet and starts being served immediately.
2) Non-preemptive priority: The server always completes the current packet and then starts serving the most recent packet in the queue.

IV. PAoI under LCFS with Preemptive Priority

We begin with LCFS with preemptive priority. Note that in this case, $u(n) = a(n), \forall n$, and $\{\hat{S}_n\}_n$ are statistically the same. As shown in Fig. 2(a), PAoI is the elapsed time from the moment when a packet $n_i \in \Psi$ arrives, until the moment when $n_{i+1} \in \Psi$ departs (recall that Ψ denotes the set of informative packets). Define the first informative packet which arrives no earlier that n as

$$\alpha(n) \triangleq \min\{n_i | n_i \in \Psi,\ a(n_i) \geq a(n)\},$$

and define $\hat{S}_n$ as the time duration from the moment n starts to receive service to the moment $\alpha(n)$ departs, i.e.,

$$\hat{S}_n \triangleq d(\alpha(n)) - u(n).$$

Note that if $n \in \Psi$, we have $\alpha(n) = n$ and $\hat{S}_n = S_n$. Moreover, define the first informative packet which arrives after n's departure as

$$\beta(n) \triangleq \min\{n_i | n_i \in \Psi,\ a(n_i) > d(n)\},$$

and the inter-departure time between n and $\beta(n)$:

$$\hat{Y}_n \triangleq d(\beta(n)) - d(n).$$

Since the packets arriving after $a(n_i)$ but before $d(n_i)$ preempt n_i and get lost upon departure (because $n_i \in \Psi$), we have (see Fig. 2(a)):

$$A_P^{LCFS,pre} = \mathbb{E}\{S_{n_i} + \hat{Y}_{n_i} | n_i \in \Psi\}. \quad (2)$$

A. Analyzing a Service Process

We use S_n to also denote the process of serving a packet n. For simplicity, we define the following symbols (notice that in other sections these symbols may have different definitions): $\tilde{p} \triangleq \mathbb{P}(\hat{S}_n \le S_n), \tilde{t} \triangleq \mathbb{E}\{\hat{S}_n | \hat{S}_n \le S_n\}, \tilde{s} \triangleq \mathbb{E}\{S_n | \hat{S}_n > S_n\}$, i.e., $\tilde{p}$ is the probability that there exists a packet that reaches the destination successfully during S_n (including n and the packets arriving after $a(n)$ but before $d(n)$) .

We first have the following lemma, based on which we will derive $\tilde{t}$ and $\tilde{s}$.

Lemma 1. *For a nonnegative random variable X, an event E and a sequence of events $E_1, E_2, \dots, E_K$ which satisfies $E_i \cap E_j = \emptyset, \forall i \ne j$ and $E = \cup_{k=1}^K E_k$, we have*

$$\mathbb{P}(E)\mathbb{E}\{X|E\} = \sum_{k=1}^{K} \mathbb{P}(E_k)\mathbb{E}\{X|E_k\}.$$

Proof. Omitted due to space limit. Please see [12]. □

The probability that $X_n \le S_n$ is $\frac{\lambda}{\lambda+\mu} = \rho_\lambda$. If that happens, the system will first serve packet $n+1$ (during which new packets may come and depart before $n+1$), then continue the service of n. Based on this observation, we have

$$\begin{aligned}\tilde{p} =& \rho_\mu p + \rho_\lambda [\mathbb{P}(\hat{S}_{n+1} \le S_{n+1}) + \mathbb{P}(\hat{S}_{n+1} > S_{n+1}) \\ & \times \mathbb{P}(\hat{S}_n \le S_n | X_n \le S_n, \hat{S}_{n+1} > S_{n+1})] \\ =& \rho_\mu p + \rho_\lambda [\tilde{p} + (1-\tilde{p})\tilde{p}], \end{aligned} \quad (3)$$

$$\begin{aligned}\tilde{p}\tilde{t} =& \rho_\mu p \mathbb{E}\{S_n | X_n > S_n, n \in \Phi\} + \rho_\lambda \big[\tilde{p}\mathbb{E}\{X_n + \\ & \hat{S}_{n+1} | X_n \le S_n, \hat{S}_{n+1} \le S_{n+1}\} + (1-\tilde{p})\tilde{p} \\ & \times \mathbb{E}\{\hat{S}_n | X_n \le S_n, \hat{S}_{n+1} > S_{n+1}, \hat{S}_n \le S_n\}\big] \\ =& \rho_\mu p \frac{1}{\lambda+\mu} + \rho_\lambda \big[\tilde{p}(\frac{1}{\lambda+\mu} + \tilde{t}) \\ & + (1-\tilde{p})\tilde{p}(\frac{1}{\lambda+\mu} + \tilde{s} + \tilde{t})\big], \end{aligned} \quad (4)$$

$$\begin{aligned}(1-\tilde{p})\tilde{s} =& \rho_\mu (1-p)\mathbb{E}\{S_n | X_n > S_n, n \notin \Phi\} + \rho_\lambda (1-\tilde{p})^2 \\ & \times \mathbb{E}\{S_n | X_n \le S_n, \hat{S}_{n+1} > S_{n+1}, \hat{S}_n > S_n\} \\ =& \rho_\mu (1-p)\frac{1}{\lambda+\mu} + \rho_\lambda (1-\tilde{p})^2 (\frac{1}{\lambda+\mu} + 2\tilde{s}). \end{aligned} \quad (5)$$

We get from (3) that:

$$\lambda \tilde{p}^2 + (\mu - \lambda)\tilde{p} - \mu p = 0, \quad (6)$$

which leads to:

$$\tilde{p} = \frac{-(\mu-\lambda) + \sqrt{(\mu-\lambda)^2 + 4\lambda\mu p}}{2\lambda}. \quad (7)$$

Solving (4) and (5), and using (6) give us:

$$(1-\tilde{p})\tilde{s} = \frac{1-\tilde{p}}{\mu - \lambda + 2\lambda\tilde{p}}, \quad (8)$$

$$\tilde{p}\tilde{t} = \frac{\tilde{p} + \frac{\mu}{\mu-\lambda+2\lambda\tilde{p}}(\tilde{p} - p)}{\mu - \lambda + \lambda\tilde{p}}. \quad (9)$$

On the other hand, $n \in \Psi$ means that only n reaches the destination successfully during S_n, or $\hat{S}_n = S_n$. Using the same method as above (see [12] for details), we can get

$$\mathbb{E}\{S_n | n \in \Psi\} = \frac{1}{\mu - \lambda + 2\lambda\tilde{p}}.$$

B. Computing PAoI

Now consider $\mathbb{E}\{\hat{Y}_{n_i} | n_i \in \Psi\} = \mathbb{E}\{\hat{Y}_{n_i}\}$. Suppose the first packet which arrives after $d(n_i)$ is $\tilde{n}_i$. Since the exponential distribution is memoryless, the expected time from $d(n_i)$ to $a(\tilde{n}_i)$ is $\frac{1}{\lambda}$. If $\hat{S}_{\tilde{n}_i} \le S_{\tilde{n}_i}$, then the (expected) remaining time of $\hat{Y}_{n_i}$ from $a(\tilde{n}_i)$ is $\tilde{t}$. Otherwise the remaining time is $\tilde{s} + \mathbb{E}\{\hat{Y}_{\tilde{n}_i}\}$. Based on the above analysis,

$$\mathbb{E}\{\hat{Y}_{n_i}\} = \frac{1}{\lambda} + \tilde{p}\tilde{t} + (1-\tilde{p})(\tilde{s} + \mathbb{E}\{\hat{Y}_{\tilde{n}_i}\}).$$

Substituting (8) and (9) into the above gives us:

$$\mathbb{E}\{\hat{Y}_{n_i}\} = \frac{\mu(\mu-\lambda) + 2\lambda\mu p + \lambda(\lambda+\mu)\tilde{p}}{\lambda\mu p(\mu - \lambda + 2\lambda\tilde{p})}.$$

As a result,

$$\begin{aligned}A_P^{LCFS,pre} =& \frac{1}{\mu-\lambda+2\lambda\tilde{p}} + \frac{\mu(\mu-\lambda) + 2\lambda\mu p + \lambda(\lambda+\mu)\tilde{p}}{\lambda\mu p(\mu-\lambda+2\lambda\tilde{p})} \\ =& \frac{\mu(\mu-\lambda) + 3\lambda\mu p + \lambda(\lambda+\mu)\tilde{p}}{\lambda\mu p(\mu-\lambda+2\lambda\tilde{p})}, \end{aligned} \quad (10)$$

where $\tilde{p}$ is given in (7). In the case when $p = 1$, the above result becomes $PAoI = \frac{1}{\lambda+\mu} + \frac{1}{\lambda} + \frac{1}{\mu}$.

V. PAoI under LCFS with Non-Preemptive Priority

In this case, if a new packet arrives while the server is busy, it cannot interrupt the current service. From Fig. 2(b), we see that PAoI is similarly the elapsed time from the moment when a packet $n_i \in \Psi$ arrives, to the moment when n_{i+1} departs. Define the first informative packet which arrives after n starts to receive service as

$$\gamma(n) \triangleq \min\{n_i | n_i \in \Psi,\ a(n_i) > u(n)\}.$$

and the time duration from the moment n starts to receive service to the moment $\gamma(n)$ departs as

$$\hat{Z}_n \triangleq d(\gamma(n)) - u(n).$$

Since the packets arriving after $a(n_i)$ but before $u(n_i)$ are served before n_i and get lost upon departure (because $n_i \in \Psi$), we have (see Fig. 2(b)):

$$A_P^{LCFS,non} = \mathbb{E}\{W_{n_i} + \hat{Z}_{n_i} | n_i \in \Psi\}. \quad (11)$$

A. Analyzing a Service Process

We first define $\bar{S}_n$ as the process since $u(n)$ till the first time the server becomes free or starts to serve a packet that arrives no later than $u(n)$ (excluding n). Since $\bar{S}_n$ is determined by the services and arrivals after $u(n)$ and independent of the system state at $u(n)$ and the history before $u(n)$, the $\bar{S}_n$ processes induced by different packets n are identically distributed, so are the $\hat{Z}_n$ processes if $\hat{Z}_n \le \bar{S}_n$. We re-define the following symbols:

$\tilde{p} \triangleq \mathbb{P}(\hat{Z}_n \le \bar{S}_n), \tilde{t} \triangleq \mathbb{E}\{\hat{Z}_n | \hat{Z}_n \le \bar{S}_n\}, \tilde{s} \triangleq \mathbb{E}\{\bar{S}_n | \hat{Z}_n > \bar{S}_n\}$, i.e., $\tilde{p}$ is the probability that there exists a packet which arrives after $u(n)$ and reaches the destination successfully during $\bar{S}_n$.

Consider $\bar{S}_n$. Suppose the number of packets arriving during S_n is $\sigma(S_n)$. We have $\forall k \ge 0$,

$$p(\sigma(S_n) = k) = (\rho_\lambda)^k \rho_\mu, \quad \mathbb{E}\{S_n | \sigma(S_n) = k\} = \tfrac{k+1}{\lambda+\mu}.$$

If $\sigma(S_n) = k > 0$ (which is needed for $\hat{Z}_n \le \bar{S}_n$), when n completes service, the system will serve the $(n+k)$-th packet and enter $\bar{S}_{n+k}$. If $n+k \in \Phi$ then $\hat{Z}_n \le \bar{S}_n$ and the remaining time of $\hat{Z}_n$ from $d(n)$ is $\mathbb{E}\{S_{n+k} | \sigma(S_n) = k, n+k \in \Phi\} = \frac{1}{\mu}$. If $n+k \notin \Phi$, then if $\hat{Z}_{n+k} \le \bar{S}_{n+k}$, we have $\hat{Z}_n \le \bar{S}_n$ and the remaining time of $\hat{Z}_n$ from $d(n)$ is $\mathbb{E}\{\hat{Z}_{n+k} | \sigma(S_n) =$

$k, n+k \notin \Phi, \hat{Z}_{n+k} \le \bar{S}_{n+k}\} = \mathbb{E}\{\hat{Z}_{n+k} | \hat{Z}_{n+k} \le \bar{S}_{n+k}\} = \tilde{t}$. Similar analyses apply to the $(n+k-1)$-th, the $(n+k-2)$-th, $\cdots$, and the $(n+1)$-th packet. Thus with Lemma 1,

$$\begin{aligned}\tilde{p} &= \sum_{k=1}^{\infty} \rho_\mu (\rho_\lambda)^k \{p + (1-p)\tilde{p} \\ &\quad +(1-p)(1-\tilde{p})[p + (1-p)\tilde{p}] + \cdots \\ &\quad +(1-p)^{k-1}(1-\tilde{p})^{k-1}[p+(1-p)\tilde{p}]\}, \end{aligned} \quad (12)$$

$$\begin{aligned}\tilde{p}\tilde{t} &= \sum_{k=1}^{\infty} \rho_\mu (\rho_\lambda)^k \Big\{ p(\frac{k+1}{\lambda+\mu} + \frac{1}{\mu}) \\ &\quad +(1-p)\tilde{p}(\frac{k+1}{\lambda+\mu} + \tilde{t}) + (1-p)(1-\tilde{p}) \\ &\quad \times \big[p(\frac{k+1}{\lambda+\mu} + \tilde{s} + \frac{1}{\mu}) + (1-p)\tilde{p}(\frac{k+1}{\lambda+\mu} \\ &\quad + \tilde{s} + \tilde{t})\big] + \cdots + (1-p)^{k-1}(1-\tilde{p})^{k-1} \\ &\quad \times \big[p(\frac{k+1}{\lambda+\mu} + k\tilde{s} - \tilde{s} + \frac{1}{\mu}) \\ &\quad + (1-p)\tilde{p}(\frac{k+1}{\lambda+\mu} + k\tilde{s} - \tilde{s} + \tilde{t})\big]\Big\}, \end{aligned} \quad (13)$$

and

$$(1-\tilde{p})\tilde{s} = \sum_{k=0}^{\infty} \rho_\mu (\rho_\lambda)^k (1-p)^k (1-\tilde{p})^k (\frac{k+1}{\lambda+\mu} + k\tilde{s}). \quad (14)$$

From (12), we get:

$$\lambda(1-p)\tilde{p}^2 + (\mu - \lambda + 2\lambda p)\tilde{p} - \lambda p = 0, \quad (15)$$

which leads to

$$\tilde{p} = \begin{cases} \frac{-(\mu-\lambda+2\lambda p) + \sqrt{(\lambda+\mu)^2 - 4\lambda\mu(1-p)}}{2\lambda(1-p)}, & 0 < p < 1 \\ \frac{\lambda}{\lambda+\mu}, & p = 1 \end{cases}. \quad (16)$$

Solving (13) and (14), and using (15) give us

$$(1-\tilde{p})\tilde{s} = \frac{1-\tilde{p}}{\lambda+\mu-2\lambda(1-p)(1-\tilde{p})}, \quad (17)$$

$$\tilde{p}\tilde{t} = \frac{\lambda p + 2\lambda p^2 + (\lambda - 2\lambda p^2 - \mu + \mu p)\tilde{p}}{\mu p[\lambda+\mu-2\lambda(1-p)(1-\tilde{p})]}. \quad (18)$$

B. Computing PAoI

Now we compute PAoI shown in Fig. 2(b). Define $\pi(t)$ as the number of packets in the system (including the packet being served) at time t. So $\pi(t) = 0$ means the system is free at time t. Different from the preemptive case, here $\pi(a(n_i))$ and $\pi(u(n_i))$ will respectively affect W_{n_i} and $\hat{Z}_{n_i}$, in that they affect the degree to which new packets need to wait for service completion.

We first compute the number of packets an arrival in Ψ sees when it arrives. Since Ψ is a special set of packets, they do not see exactly as what an ordinary packet will see. To this end, we define for each k

$$p_k \triangleq \mathbb{P}[\pi(a(n)) = k | n \in \Psi] = \frac{\mathbb{P}[\pi(a(n)) = k, n \in \Psi]}{\mathbb{P}(n \in \Psi)}. \quad (19)$$

Consider the waiting time W_n of packet n. If $\pi(a(n)) = 0$ then $W_n = 0$. Otherwise n needs to wait for the completion of the current service and the services of packets which arrive during the current service, till the server starts to serve a packet arriving no later that $a(n)$. Since the exponential distribution is memoryless, for $\pi(a(n)) > 0$, W_n is the same as the process $\bar{S}_{\bar{n}}$ of a virtual packet $\bar{n}$ with $u(\bar{n}) = a(n)$, and $n \in \Psi$ is equivalent to $(\hat{Z}_{\bar{n}} > \bar{S}_{\bar{n}}) \cap (n \in \Phi)$. For a steady-state $M/M/1$ queue, we know that $\mathbb{P}[\pi(t) = k] = (1-\rho)(\rho)^k$. Thus,

$$p_0 = \frac{(1-\rho)p}{\mathbb{P}[n \in \Psi]}, \quad p_k = \frac{(1-\rho)(\rho)^k(1-\tilde{p})p}{\mathbb{P}[n \in \Psi]}, k \ge 1.$$

Moreover, $\sum_{k=0}^{\infty} p_k$=1. Therefore,

$$p_0 = \frac{\mu - \lambda}{\mu - \lambda\tilde{p}}, \quad p_k = \frac{\mu-\lambda}{\mu-\lambda\tilde{p}}(1-\tilde{p})(\rho)^k, k \ge 1.$$

Hence, the waiting time can be computed as:

$$\begin{aligned}\mathbb{E}\{W_{n_i} | n_i \in \Psi\} &= p_0 \cdot 0 + (1-p_0)\tilde{s} \\ &= \frac{\lambda(1-\tilde{p})}{(\mu-\lambda\tilde{p})[\lambda+\mu-2\lambda(1-p)(1-\tilde{p})]}. \end{aligned}$$

For $\mathbb{E}\{\hat{Z}_{n_i} | n_i \in \Psi\}$, define

$$z_k \triangleq \mathbb{E}\{\hat{Z}_n | \pi(u(n)) = k, n \in \Psi\} = \mathbb{E}\{\hat{Z}_n | \pi(u(n)) = k\}.$$

For $\mathbb{E}\{\hat{Z}_n | \pi(u(n)) = k\}$, if a packet n_j arrives during $S(n)$ (with probability $\frac{\lambda}{\lambda+\mu} = \rho_\lambda$), it will wait $W_{n_j} = \bar{S}_{\bar{n}_j}$ before being served, with $\bar{n}_j$ a virtual packet defined as before. Since $\pi(u(n_j)) = k$, if $\hat{Z}_{\bar{n}_j} > \bar{S}_{\bar{n}_j}$ and $n_j \notin \Phi$, the (expected) remaining time of $\hat{Z}_n$ from $u(n_j)$ is still z_k. Otherwise no packet arrives during $S(n)$, giving us $\pi(d(n)) = k-1$. Based on the above analysis, we get:

$$\begin{aligned} z_1 &= \rho_\mu \big[\frac{1}{\lambda+\mu} + \frac{1}{\lambda} + p\frac{1}{\mu} + (1-p)z_1\big] + \rho_\lambda \big[\frac{1}{\lambda+\mu} \\ &\quad + \tilde{p}\tilde{t} + (1-\tilde{p})p(\tilde{s} + \frac{1}{\mu}) + (1-\tilde{p})(1-p)(\tilde{s}+z_1)\big], \end{aligned} \quad (20)$$

and that for general k,

$$\begin{aligned} z_k &= \rho_\mu (\frac{1}{\lambda+\mu} + z_{k-1}) + \rho_\lambda \big[\frac{1}{\lambda+\mu} \\ &\quad + \tilde{p}\tilde{t} + (1-\tilde{p})p(\tilde{s} + \frac{1}{\mu}) + (1-\tilde{p})(1-p)(\tilde{s}+z_k)\big]. \end{aligned} \quad (21)$$

Solving (20) gives us

$$z_1 = \frac{\mu + \lambda + \lambda p + \lambda^2\tau}{\lambda[\lambda + \mu p - \lambda(1-p)(1-\tilde{p})]}, \quad (22)$$

where

$$\tau = \frac{(\lambda+\mu)p + (\lambda+\mu)p^2 + [\lambda + (\mu-\lambda)p^2 - \mu]\tilde{p}}{\mu p[\lambda+\mu-2\lambda(1-p)(1-\tilde{p})]}. \quad (23)$$

From (21), we can get

$$z_k - \frac{(1-\tilde{p})(1+\lambda\tau)}{\mu\tilde{p}} = (1-\tilde{p})[z_{k-1} - \frac{(1-\tilde{p})(1+\lambda\tau)}{\mu\tilde{p}}].$$

The evolution of the LCFS queueing system shows that if a packet n sees no more than one packet when it arrives, then there will be only one packet in the system (packet n itself) when it starts to receive service. Thus,

$$\mathbb{P}[\pi(u(n)) = 1 | n \in \Psi] = p_0 + p_1, \quad (24)$$

$$\mathbb{P}[\pi(u(n)) = k | n \in \Psi] = p_k, k \ge 2. \quad (25)$$

Hence,

$$\begin{aligned} \mathbb{E}\{\hat{Z}_{n_i} | n_i \in \Psi\} &= (p_0 + p_1)z_1 + \sum_{k=2}^{\infty} p_k z_k \\ &= p_0 z_1 + \sum_{k=1}^{\infty} p_k (1-\tilde{p})^{k-1}[z_1 - \frac{(1-\tilde{p})(1+\lambda\tau)}{\mu\tilde{p}}] \\ &\quad + \sum_{k=1}^{\infty} p_k \frac{(1-\tilde{p})(1+\lambda\tau)}{\mu\tilde{p}} \end{aligned}$$

Substituting (22), PAoI can be computed as:

$$A_P^{LCFS,non} = \frac{\lambda(1-\tilde{p})}{(\mu-\lambda\tilde{p})[\lambda+\mu-2\lambda(1-p)(1-\tilde{p})]}$$

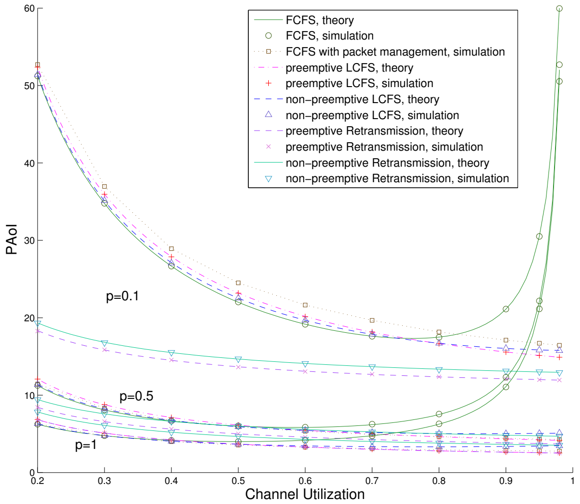


Fig. 3. PAoI in different queueing systems with packet loss.

$$+\frac{\mu(\mu-\lambda)(\mu+\lambda+\lambda p+\lambda^2\tau)}{\lambda(\mu-\lambda\tilde{p})[\mu-\lambda(1-\tilde{p})][\lambda+\mu p-\lambda(1-p)(1-\tilde{p})]}$$
$$+\frac{\lambda^2(1-\tilde{p})^2(1+\lambda\tau)}{\mu(\mu-\lambda\tilde{p})[\mu-\lambda(1-\tilde{p})]}, \tag{26}$$

where τ is given in (23) and $\tilde{p}$ is given in (16).

VI. PAoI under Retransmission with Preemptive Priority or Non-Preemptive Priority

The results are (For rigorous derivation, please see [12]):

$$A_P^{RT,pre} = \frac{1}{\lambda+p\mu}+\frac{1}{\lambda}+\frac{1}{p\mu}; \tag{27}$$

$$A_P^{RT,non} = \frac{1}{\mu}+\frac{1}{\lambda+p\mu}+\frac{1}{\lambda}+\frac{1}{p\mu}. \tag{28}$$

Remark: It turns out that result (27) corresponds to the result under the LCFS with preemptive priority policy with a service rate $p\mu$ and a success probability 1. This is intuitive since in the LCFS with preemptive priority case with $p = 1$, each packet is either transmitted successfully or preempted, while in this case each packet is still either transmitted successfully or preempted, with a mean service time $\frac{1}{p\mu}$.

VII. Numerical results

We present numerical evaluations of PAoI under different scheduling policies, including FCFS, FCFS with packet management (the $M/M/1/2^*$ scheme in [8]), LCFS with preemptive priority, LCFS with non-preemptive priority, Retransmission with preemptive priority and Retransmission with non-preemptive priority. Note that the $M/M/1/2^*$ scheme in [8] is equivalent to the LCFS with non-preemptive priority policy that discards all stale packets. The service rate is set to $\mu = 1$ while the arrival rate is varied to show performances under different channel utilizations $\rho = \frac{\lambda}{\mu}$. The cases $p = 0.1$, $p = 0.5$ and $p = 1$ are selected to represent different delivery error regimes. We present not only the results computed from our formulas (1), (10), (26), (27) and (28), but also those obtained by simulating real queueing systems with the corresponding settings.

From Fig. 3, we see that the simulation results match our theoretical results very well. We can see that when channel utilization is high, the PAoI under FCFS becomes very large due to large queueing delay, while other policies effectively avoid this problem. On the other hand, when packet loss rate is high, FCFS with packet management suffers from the lack of packet deliveries but LCFS again ensures a low PAoI, matching our intuition about the benefits of LCFS. Moreover, retransmission policies have significant reductions on PAoI compared to other policies when packet loss rate is high. But when packet loss rate is low, Retransmission with non-preemptive priority suffers a performance loss since retransmissions can also block later packets.

VIII. Conclusion

We consider the peak age-of-information (PAoI) in an $M/M/1$ queueing system with packet delivery failure, a setting that models real-world situations with transmission errors. We derive exact PAoI expressions under different scheduling policies, including FCFS, LCFS with preemptive priority, LCFS with non-preemptive priority, Retransmission with preemptive priority, and Retransmission with non-preemptive priority. Our analytical and simulation results show that the LCFS principle as well as retransmissions can successfully avoid increments in PAoI resulting from large queueing delay and packet loss.

Acknowledgment

The authors would like to thank Prof. Eytan Modiano at MIT for the motivating discussions and valuable comments.

This work was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61361136003, 61303195, Tsinghua Initiative Research Grant, Microsoft Research Asia Collaborative Research Award, and the China Youth 1000-talent Grant.

References

[1] P. Corke, T. Wark, R. Jurdak, et al. Environmental wireless sensor networks. *Proceedings of the IEEE*, 2010, 98(11): 1903-1917.

[2] P. Papadimitratos, A. La Fortelle, K. Evenssen, et al. Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation. *Communications Magazine*, IEEE, 2009, 47(11): 84-95.

[3] A. Reddy, S. Banerjee, A. Gopalan, et al. On distributed scheduling with heterogeneously delayed network-state information. *Queueing Systems*, 2012, 72(3-4): 193-218.

[4] S. Kaul, R. Yates, and M. Gruteser. Real-time status: How often should one update? *Proceedings of IEEE International Conference on Computer Communications, INFOCOM*, 2012.

[5] R. D. Yates and S. Kaul. Real-time status updating: Multiple sources. *Proceedings of IEEE ISIT*, 2012.

[6] S. K. Kaul, R. D. Yates, and M. Gruteser. Status Updates Through Queues. *Proceedings of IEEE Annual Conference on Information Sciences and Systems, CISS*, 2012.

[7] C. Kam, S. Kompella, and A. Ephremides. Age of information under random updates. *Proceedings of IEEE ISIT*, 2013.

[8] M. Costa, M. Codreanu, and A. Ephremides. Age of information with packet management. *Proceedings of IEEE ISIT*, 2014.

[9] L. Huang and E. Modiano. Optimizing Age-of-Information in a Multi-class Queueing System. *Proceedings of IEEE ISIT*, 2015.

[10] X. Guo, R. Singh, P.R. Kumar, et al. A High Reliability Asymptotic Approach for Packet Inter-Delivery Time Optimization in Cyber-Physical Systems. *Proceedings of ACM MobiHoc*, 2015.

[11] Y. Sun, E. Uysal-Biyikoglu, R. Yates, et al. Update or Wait: How to Keep Your Data Fresh. *arXiv preprint arXiv:1601.02284*, 2016.

[12] K. Chen and L. Huang. Age-of-Information in the Presence of Error. *arXiv techreport arXiv:1605.00559*, 2016.