A Bound on Peak Age of Information Distribution

Zhidu Li[1,2,3], Ailing Zhong[1,2,3], Yuming Jiang[4], Tong Tang[1,2,3], Ruyan Wang[1,2,3]
[1] School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, China
[2] Advanced Network and Intelligent Connection Technology Key Laboratory of Chongqing Education Commission of China
[3] Chongqing Key Laboratory of Ubiquitous Sensing and Networking, China
[4] NTNU – Norwegian University of Science and Technology, Norway
Email: lizd@cqupt.edu.cn

***Abstract*—This paper presents a study on peak age of information (AoI), focusing on its distribution that is more important for AoI guarantees than the mean. Specifically, the relation of peak AoI to the underlying information generation and transmission processes is explicitly formulated. Based on this formulation and by exploring the independence information between the information generation and transmission processes, a general bound on the distribution of peak AoI is derived. To showcase the use of the derived bound, it is applied to two representative cases, which are characterized by the M/M/1 and D/M/1 queuing models. Numerical results obtained from the proposed bound analysis are finally introduced and discussed in comparison with exact results, validating the bound.**

***Index Terms*—Peak age of information, performance guarantee, probabilistic bound, martingale theory.**

I. Introduction

The timeliness or the freshness of information is of great significance for real-time sensing applications, such as factory automation, Metaverse, etc. [1]. It is related not only to the end-to-end delay of information but also to the relevance of the information generated. Traditional performance metrics such as delay or latency are unable to characterize the information freshness [2]. In this regard, age of information (AoI) and peak AoI are introduced and have become prevailing metrics to quantify the freshness of information [3]. Particularly, the peak AoI is usually applied to characterize the instantaneous information freshness when an update information is received by the destination node [4].

As information may be generated and transmitted through wireless networks, the stochastic properties of wireless channels may have great influence on the peak AoI performance. Besides, due to the application requirement, the information generation process may also be stochastic. In such cases, it is not possible to provide deterministic peak AoI guarantee to the users. Consequently, it is necessary to study how to guarantee the peak AoI performance from the probabilistic point of view.

In the literature, most contributions have been devoted to average or expected (peak) AoI analysis and guaranteeing. In [5], a scheduling policy between associated devices was proposed to reduce the average AoI in Internet of Things. In [6], service preemption was introduced to improve the average peak AoI for a generate-at-will source node. In [7], an asymptotic expression of the average peak AoI was derived to determine the superiority of overlay and underlay schemes in a primary IoT system. In [8], the focus was on the peak AoI performance guarantee for massive machine type communication devices, where a closed-form expression of expected peak AoI was derived by taking the energy harvesting process into account.

To further explore the statistical characteristics of information freshness, some researchers tried to find out the distribution of peak AoI using queueing models [9]. In [10], closed-form expressions of peak AoI distribution were derived for the M/M/1 and M/D/1 queueing models. In [11], the distribution of peak AoI was studied based on three different queueing disciplines under the assumption of Bernoulli update arrivals. The work [12] analyzed the peak AoI violation probability with Mellin transform technique in a UAV communication network where the data packets are generated according to a Bernoulli process. In [2], the statistical bounds of AoI was studied for the worst case based on min-plus theory. However, as the independence information of the arrival and service processes, such as independent and identically distributed (i.i.d) increments in them, is not fully exploited in [12] and [2], the obtained probabilistic peak AoI bound is loose and non-universal.

The objective of this work is to analyze the peak AoI distribution under a general setting and derive a (tight) bound. To this aim, based on sample-path analysis, we first investigate the relation of peak AoI with the information generation and transmission processes. Specifically, it is found that the peak AoI immediate before the arrival of a new update can be formulated with the inter-arrival times and transmission times of the early adjacent update. Based on this result, only assuming they are respectively independent and identically distributed (i.i.d.), a general bound on the peak AoI distribution is derived. As a highlight, the martingale theory is applied, which is a known technique to tighten the obtained bounds. Thereafter, the application of the proposed analytical approach and obtained bound are exemplified with two specific yet representative cases, where upper bounds on peak AoI violation probability and on average peak AoI are derived. Finally, numerical results and comparison with the exact results obtained from queueing theory analysis are presented and discussed. The comparison not only validates but also implies the tightness of the bounds.

The remaining of this paper is organized as follows. Section II derives the general probabilistic bound of the peak AoI. Case analysis under specific scenarios is presented in Section III.

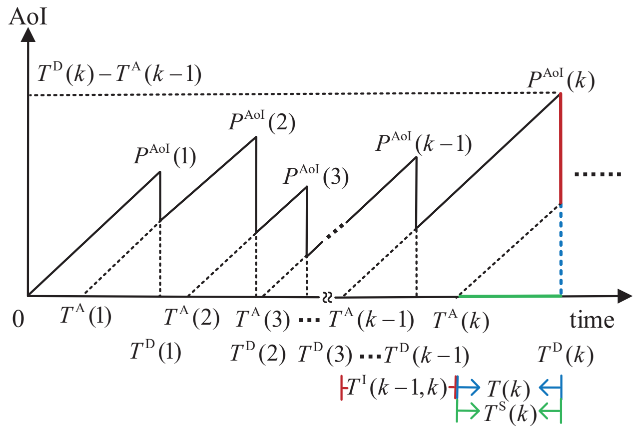

Fig. 1. Evolution of the AoI for information updates.

In Section IV, numerical results are provided and discussed. Section V concludes the paper.

II. Bound on Peak AoI Distribution

In this section, a general bound on the distribution of peak AoI is derived. Throughout this paper, we use $f_X(\cdot)$ and $F_X(\cdot)$ to represent the probability density function (PDF) and the cumulative distribution function (CDF) of random variable X, respectively. Also, let $\mathrm{E}[\cdot]$ denote the expectation function.

A. Relation of Peak AoI to Inter-Information Update Time and Information Transmission Time

As depicted in Fig. 1, the AoI curve reaches a peak immediately before an information update is received by the destination node. Specifically, the peak AoI corresponding to the kth update, denoted as $P^{\mathrm{AoI}}(k)$, can be written as

$$P^{\mathrm{AoI}}(k) = T^{\mathrm{D}}(k) - T^{\mathrm{A}}(k-1) \tag{1}$$

where $T^{\mathrm{A}}(k)$ and $T^{\mathrm{D}}(k)$ denote the arrival time and the departure time of the kth update, respectively. Without loss of generality, we set $T^{\mathrm{A}}(0) = 0$ [12].

Under the first-come-first-served (FCFS) policy, the departure time of the kth update can be expressed as

$$T^{\mathrm{D}}(k) = \max_{1\le j\le k}\{T^{\mathrm{A}}(j) + T^{\mathrm{SC}}(j,k)\} \tag{2}$$

where $T^{\mathrm{SC}}(j,k)$ denotes the cumulative transmission time from the jth update to the kth update ($j < k$) [13], and there holds

$$T^{\mathrm{SC}}(j,k) = \sum\nolimits_{n=j}^{k} T^{\mathrm{S}}(n) \tag{3}$$

Here, $T^{\mathrm{S}}(n)$ denotes the transmission time of the nth update. It is worth mentioning that the total sojourn time $T(k)$ for any update is not less than $T^{\mathrm{S}}(k)$.

Similarly, let $T^{\mathrm{I}}(k-1,k) = T^{\mathrm{A}}(k) - T^{\mathrm{A}}(k-1)$ denote the inter-arrival time between the $(k-1)$th update and the kth update for $k = 1, 2, \cdots$. Then, the inter-arrival time between the jth update and the kth update can be written as

$$T^{\mathrm{I}}(j,k) = \sum\nolimits_{n=j+1}^{k} T^{\mathrm{I}}(n-1,n) \tag{4}$$

We assume that the inter-arrival time $T^{\mathrm{I}}(k-1,k)$ and the transmission time $T^{\mathrm{S}}(k)$ are both i.i.d for each update, respectively. In addition we assume the following stability condition:

$$\mathrm{E}[T^{\mathrm{I}}(k-1,k)] \ge \mathrm{E}[T^{\mathrm{S}}(k)] \tag{5}$$

By integrating Eqs. (2), (3) and (4), the peak AoI in Eq. (1) can be further expressed as

$$\begin{aligned}
P^{\mathrm{AoI}}(k) &= \max_{1\le j\le k}\{T^{\mathrm{A}}(j) + T^{\mathrm{SC}}(j,k) - T^{\mathrm{A}}(k-1)\} \\
&= \max\{\max_{1\le j\le k-1}\{T^{\mathrm{A}}(j) - T^{\mathrm{A}}(k-1) + T^{\mathrm{SC}}(j,k)\}, \\
&\qquad T^{\mathrm{A}}(k) - T^{\mathrm{A}}(k-1) + T^{\mathrm{S}}(k)\} \\
&= \max\{\max_{1\le j\le k-1}\{T^{\mathrm{SC}}(j,k) - T^{\mathrm{I}}(j,k-1)\}, \\
&\qquad T^{\mathrm{I}}(k-1,k) + T^{\mathrm{S}}(k)\} \\
&= \max\{\max_{1\le j\le k-1}\{T^{\mathrm{SC}}(j,k-1) - T^{\mathrm{I}}(j,k-1)\} + T^{\mathrm{S}}(k), \\
&\qquad T^{\mathrm{I}}(k-1,k) + T^{\mathrm{S}}(k)\} \\
&= \max\{\max\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2} T^{\mathrm{S}}(n) - \sum_{n=j+1}^{k-1} T^{\mathrm{I}}(n-1,n)), 0\} \\
&\qquad + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} + T^{\mathrm{S}}(k) \\
&= \max\{\max\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2} [T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]), 0\} \\
&\qquad + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} + T^{\mathrm{S}}(k)
\end{aligned} \tag{6}$$

From Eq. (6), it is worth noting that the peak AoI is affected by not only the inter-arrival time $T^{\mathrm{I}}(k-1,k)$ but also the transmission time $T^{\mathrm{S}}(k)$. Hence, it is impossible to provide deterministic peak AoI guarantee if $T^{\mathrm{I}}(k-1,k)$ or $T^{\mathrm{S}}(k)$ is random. In what follows, the peak AoI performance is analyzed from a probabilistic bound or violation probability point of view.

B. Peak AoI Violation Probability

We denote an auxiliary parameter B, where $B = \max\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]), 0\}$. Then, Eq. (6) can be further simplified as

$$P^{\mathrm{AoI}}(k) = \max\{B + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} + T^{\mathrm{S}}(k) \tag{7}$$

Since $T^{\mathrm{I}}(k-1,k)$ and $T^{\mathrm{S}}(k)$ are both i.i.d parameters for $\forall k = 1, 2, \cdots$, for a given peak AoI threshold d, the corresponding violation probability can be expressed as

$$\begin{aligned}
&\Pr\{P^{\mathrm{AoI}}(k) > d\} \\
&= \Pr\{\max\{B + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} + T^{\mathrm{S}}(k) > d\} \\
&= 1 - \Pr\{\max\{B + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \\
&\qquad + T^{\mathrm{S}}(k) \le d\} \\
&= 1 - (\int_0^d \{\Pr\{\max\{B + T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\} \\
&\qquad \cdot f_{T^{\mathrm{S}}}(d-x)\}dx)
\end{aligned} \tag{8}$$

In order to further reveal the probabilistic characteristics of the peak AoI, a lemma is introduced in the following.

Lemma 1: For $\forall k = 1, 2, \cdots$, there always holds:

$$\begin{aligned}\Pr\{\max\{B+T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\}\\ \ge \Pr\{B+T^{\mathrm{I}}(k-1,k)\} \le x\}\end{aligned} \tag{9}$$

Proof: We prove this lemma through taking the following two cases into account.

Case 1: If $B+T^{\mathrm{S}}(k-1) \ge T^{\mathrm{I}}(k-1,k)$, we have

$$\begin{aligned}&\Pr\{\max\{B+T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\}\\ &= \Pr\{B+T^{\mathrm{S}}(k-1) \le x\}\\ &= \int_0^x F_{T^{\mathrm{S}}}(y) f_B(x-y) dy\\ &\overset{(a)}{\ge} \int_0^x F_{T^{\mathrm{I}}}(y) f_B(x-y) dy\\ &= \Pr\{B+T^{\mathrm{I}}(k-1,k) \le x\}\end{aligned} \tag{10}$$

Here, step (a) holds because $\int_0^\infty (1-F_{T^{\mathrm{I}}}(y))dy = \mathrm{E}[T^{\mathrm{I}}(k-1,k)] \ge \mathrm{E}[T^{\mathrm{S}}(k)] = \int_0^\infty (1-F_{T^{\mathrm{S}}}(y))dy$, and both $F_{T^{\mathrm{S}}}(y)$ and $F_{T^{\mathrm{I}}}(y)$ are increasing functions, where, with the stability condition, $\mathrm{E}[T^{\mathrm{I}}(k-1,k)] \ge \mathrm{E}[T^{\mathrm{S}}(k)]$, meaning the average inter-arrival time is not less than the average transmission time.

Case 2: If $B+T^{\mathrm{S}}(k-1) < T^{\mathrm{I}}(k-1,k)$, since $B \ge 0$ by definition, we have

$$\begin{aligned}&\Pr\{\max\{B+T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\}\\ &= \Pr\{T^{\mathrm{I}}(k-1,k) \le x\}\\ &\ge \Pr\{B+T^{\mathrm{I}}(k-1,k) \le x\}\end{aligned} \tag{11}$$

Thus, $\Pr\{\max\{B+T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\} \ge \Pr\{B+T^{\mathrm{I}}(k-1,k) \le x\}$ always holds, which completes the proof. ■

Applying Lemma 1 in Eq. (8), the peak AoI violation probability can be further derived as

$$\begin{aligned}&\Pr\{P^{\mathrm{AoI}}(k) > d\}\\ &= 1 - (\int_0^d \{\Pr\{\max\{B+T^{\mathrm{S}}(k-1), T^{\mathrm{I}}(k-1,k)\} \le x\}\\ &\qquad \cdot f_{T^{\mathrm{S}}}(d-x)\}dx)\\ &\le 1 - \int_0^d \Pr\{B+T^{\mathrm{I}}(k-1,k) \le x\} f_{T^{\mathrm{S}}}(d-x)dx\\ &\overset{a}{=} 1 - \int_0^d \int_0^x \Pr\{B \le y\} f_{T^{\mathrm{I}}}(x-y) dy f_{T^{\mathrm{S}}}(d-x)dx\end{aligned} \tag{12}$$

Here, step (a) holds since B is independent with $T^{\mathrm{I}}(k-1,k)$. For a given information update policy, the characteristic information of inter-arrival time and that of transmission time of the updates are available. Then, an upper bound on the peak AoI violation probability is readily found if the probabilistic characteristics of the auxiliary parameter B can be derived. Specifically, a probabilistic upper bound of B is derived and summarized as the following lemma.

Lemma 2: For two independent sets of random variables $\{T^{\mathrm{I}}(n-1,n)\}$ and $\{T^{\mathrm{S}}(n)\}$, where $n = 1, 2, \cdots$, in each set, all the random variables are i.i.d. Let the auxiliary parameter $B = \max\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]), 0\}$. If $\mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[^{-\theta T^{\mathrm{I}}(0,1)}] \le 1$, there holds

$$\Pr\{B \le y\} \ge 1 - e^{-\theta y} \tag{13}$$

for all $\theta \ge 0$ and $y > 0$.

Proof: For any $y > 0$, we have

$$\begin{aligned}&\Pr\{B > y\}\\ &= 1 - \Pr\{B \le y\}\\ &= 1 - (\Pr\{\max\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]),\\ &\qquad 0\} \le y\})\\ &\le 1 - (\Pr\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]) \le y\}\\ &\qquad \cdot \Pr\{0 \le y\})\\ &= 1 - \Pr\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]) \le y\}\\ &= \Pr\{\max_{1\le j\le k-2}(\sum_{n=j}^{k-2}[T^{\mathrm{S}}(n) - T^{\mathrm{I}}(n,n+1)]) > y\}\end{aligned} \tag{14}$$

Then, let $V_j = e^{\theta(T^{\mathrm{S}}(k-j-2,k-2) - T^{\mathrm{I}}(k-j-3,k-2))}$, $Z_n = T^{\mathrm{S}}(n)$ and $Y_n = T^{\mathrm{I}}(n-1,n)$ i.e., $V_j = e^{\theta(Z_{k-j-2} - Y_{k-j-2} + Z_{k-j-1} - Y_{k-j-1} + \cdots + Z_{k-2} - Y_{k-2})}$. There holds:

$$\begin{aligned}V_{j+1} &= e^{\theta(T^{\mathrm{S}}(k-j-3,k-2) - T^{\mathrm{I}}(k-j-4,k-2))}\\ &= e^{\theta(Z_{k-j-3} - Y_{k-j-3} + Z_{k-j-2} - Y_{k-j-2} + \cdots + Z_{k-2} - Y_{k-2})}\\ &= V_j e^{\theta(Z_{k-j-3} - Y_{k-j-3})}\end{aligned} \tag{15}$$

Since Y_k and Z_k both have i.i.d increments, then we have

$$\begin{aligned}&\mathrm{E}[V_{j+1}|V_1, V_2, \cdots, V_j]\\ &= \mathrm{E}[V_j e^{\theta(Z_{k-j-3} - Y_{k-j-3})}|Z_{k-2}, Z_{k-3}, \cdots, Z_{k-j-2},\\ &\qquad Y_{k-2}, Y_{k-3}, \cdots, Y_{k-j-2}]\\ &\overset{a}{=} \mathrm{E}[V_j|Z_{k-2}, Z_{k-3}, \cdots, Z_{k-j-2}, Y_{k-2}, Y_{k-3}, \cdots, Y_{k-j-2}]\\ &\qquad \mathrm{E}[e^{\theta Z_{k-j-3}}]\mathrm{E}[^{-\theta Y_{k-j-3}}]\\ &\overset{b}{=} V_j \mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[^{-\theta T^{\mathrm{I}}(0,1)}]\\ &\overset{c}{\le} V_j\end{aligned} \tag{16}$$

Here, step (a) holds because Z_{k-j-3} and Y_{k-j-3} are independent each other, and also independent of $\{Z_{k-2}, Z_{k-3}, \cdots, Z_{k-j-2}, Y_{k-2}, Y_{k-3}, \cdots, Y_{k-j-2}\}$. Step (b) holds because process Y_k and Z_k both have identical increments, i.e., for the random transmission time, we have

$$\mathrm{E}[e^{\theta Z_{k-j-3}}] = \mathrm{E}[e^{\theta T^{\mathrm{S}}(k-j-3)}] = \mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}] \tag{17}$$

Correspondingly, for the random inter-arrival time, there holds:

$$\mathrm{E}[e^{-\theta Y_{k-j-3}}] = \mathrm{E}[e^{-\theta T^{\mathrm{I}}(k-j-4,k-j-3)}] = \mathrm{E}[e^{-\theta T^{\mathrm{I}}(0,1)}] \tag{18}$$

In addition, step (c) holds when $\mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[^{-\theta T^{\mathrm{I}}(0,1)}] \leq 1$.

Hence, $V_1, V_2, V_3, \cdots, V_j, \cdots, V_{k-2}$ form a non-negative supermartingale. We further have

$$\begin{aligned}\Pr\{B > y\} &\leq \Pr\{\max_{1\leq j\leq k-2}(\sum_{n=j}^{k-2}[Z_n - Y_{n+1}]) > y\} \\ &= \Pr\{\max_{1\leq j\leq k-2}\{V_{k-j-3}\} > e^{\theta y}\} \\ &\overset{a}{=} \Pr\{V_{j^*} > e^{\theta y}\} \\ &\overset{b}{\leq} e^{-\theta y}\mathrm{E}[V_{j^*}] \\ &\overset{c}{\leq} e^{-\theta y}\mathrm{E}[V_1] \\ &= e^{-\theta y}\mathrm{E}[e^{\theta Z_{k-j-3}}]\mathrm{E}[e^{-\theta Y_{k-j-3}}] \\ &\overset{d}{\leq} e^{-\theta y}\end{aligned} \tag{19}$$

Here, in step (a), $V_{j^*}(1 \leq j^* \leq k-2)$ represents the maximal value among $\{V_1, V_2, V_3, \cdots, V_j, \cdots, V_{k-2}\}$. Using the Chernoff's inequality, we can complete the derivation of step (b). In step (c), as $V_1, V_2, V_3, \cdots, V_j, \cdots, V_{k-2}$ form a non-negative supermartingale, and based on Doob's inequality for submartingales and the formulation for supermartingales, there holds $\mathrm{E}[V^*] \leq \mathrm{E}[V_1]$ [14]. In addition, the definition of V_1 and the independence between $T^{\mathrm{S}}(k)$ and $T^{\mathrm{I}}(k-1,k)$ are contributed to step (d).

Hence, there holds

$$\Pr\{B \leq y\} \geq 1 - e^{-\theta y} \tag{20}$$

which completes the proof. ■

Applying Lemma 2 to Eq. (12), the boundary of peak AoI violation probability finally holds as

$$\begin{aligned}&\Pr\{P^{\mathrm{AoI}}(k) > d\} \\ &\leq 1 - \int_0^d \int_0^x \Pr\{B \leq y\} f_{T^{\mathrm{I}}}(x-y)dy f_{T^{\mathrm{S}}}(d-x)dx \\ &\leq 1 - \int_0^d \int_0^x (1 - e^{-\theta y}) f_{T^{\mathrm{I}}}(x-y)dy f_{T^{\mathrm{S}}}(d-x)dx \\ &= 1 - \int_0^d (1 - e^{-\theta y})[f_{T^{\mathrm{I}}} * f_{T^{\mathrm{S}}}(d-y)]dy\end{aligned} \tag{21}$$

where $*$ denotes the convolution operator.

Note that in Eq. (21), θ is a non-negative parameter meeting $\mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[^{-\theta T^{\mathrm{I}}(0,1)}] \leq 1$. In addition, it is easily verified that $\Pr\{P^{\mathrm{AoI}}(k) > d\}$ is a decreasing function in θ. Hence, the probabilistic bound of peak AoI can be tightened when θ is chosen according to the following expression

$$\theta^* = \max\{\theta : \mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[^{-\theta T^{\mathrm{I}}(0,1)}] \leq 1\} \tag{22}$$

Also note that as the optimal θ is also related to the characteristics of the inter-arrival time and that of the transmission time, according to Eq. (22), the result of peak AoI violation probability in Eq. (21) can be applied to any scenario as long as the characteristics of the inter-arrival time and that of the transmission time are available. It is highlighted that those characteristics can be obtained under a given information update policy.

III. Case Study

In this section, the application of the derived result in Eq. (21) is illustrated with the help of two classical queueing models, i.e., the M/M/1 queueing model and the D/M/1 queueing model.

A. *M/M/1 Queuing Model*

In M/M/1 queuing model, the inter-arrival time and transmission time of each update are i.i.d. Let λ denote the average inter-arrival time between any two adjacent updates and μ denote the average transmission time of an update[1], the PDF of those two parameters holds as

$$f_{T^{\mathrm{I}}}(y) = \frac{1}{\lambda}e^{-\frac{1}{\lambda}y}, \quad f_{T^{\mathrm{S}}}(x) = \frac{1}{\mu}e^{-\frac{1}{\mu}x} \tag{23}$$

According to Eq. (21), the peak AoI violation probability of M/M/1 queueing model holds as

$$\begin{aligned}&\Pr\{P^{\mathrm{AoI}}(k) > d\} \\ &\leq 1 - \int_0^d (1 - e^{-\theta y})[f_{T^{\mathrm{I}}} * f_{T^{\mathrm{S}}}(d-y)]dy \\ &= 1 - \int_0^d (1 - e^{-\theta(d-y)})[f_{T^{\mathrm{I}}} * f_{T^{\mathrm{S}}}(y)]dy \\ &= 1 - \frac{1}{\mu-\lambda}\int_0^d (1 - e^{-\theta(d-y)})(e^{-\frac{1}{\mu}y} - e^{-\frac{1}{\lambda}y})dy \\ &= \frac{\mu^2\theta}{(\lambda-\mu)(1-\mu\theta)}e^{-\frac{1}{\mu}d} - \frac{\lambda^2\theta}{(\lambda-\mu)(1-\lambda\theta)}e^{-\frac{1}{\lambda}d} \\ &\qquad + \frac{1}{(1-\mu\theta)(1-\lambda\theta)}e^{-\theta d}\end{aligned} \tag{24}$$

Here, θ can be optimized according to Eq. (22), there holds

$$\begin{aligned}\mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[e^{-\theta T^{\mathrm{I}}(0,1)}] &= \frac{1}{1-\mu\theta}\cdot\frac{1}{1+\lambda\theta} \leq 1 \\ &\Rightarrow \theta^* = \frac{\lambda-\mu}{\lambda\mu}\end{aligned} \tag{25}$$

Substituting Eq. (25) into Eq. (24), we further have

$$\Pr\{P^{\mathrm{AoI}}(k) > d\} \leq \frac{\lambda}{2\mu-\lambda}e^{-\frac{(\lambda-\mu)d}{\lambda\mu}} + e^{-\frac{d}{\mu}} - \frac{\lambda}{2\mu-\lambda}e^{-\frac{d}{\lambda}} \tag{26}$$

As a result, an upper bound on the average peak AoI for the case of M/M/1 is obtained as

$$\begin{aligned}&\mathrm{E}[P^{\mathrm{AoI}}(k)] = \int_0^\infty \Pr\{P^{\mathrm{AoI}}(k) > d\}dd \\ &\leq \int_0^\infty \frac{\lambda}{2\mu-\lambda}e^{-\frac{(\lambda-\mu)d}{\lambda\mu}} + e^{-\frac{d}{\mu}} - \frac{\lambda}{2\mu-\lambda}e^{-\frac{d}{\lambda}}dd \\ &= \frac{\lambda^2}{(\lambda-\mu)} + \mu\end{aligned} \tag{27}$$

[1]Note: Here λ and μ are defined to the time while not the rate.

B. D/M/1 Queuing Model

In D/M/1 queueing model, the inter-arrival time between two adjacent updates are deterministic while the transmission time of each update follows the identically exponential distribution. Let D denote the inter-arrival time and μ denote the average transmission time of an update. When $d > D$, the peak AoI violation probability of D/M/1 queueing model holds as

$$\begin{aligned}
&\Pr\{P^{\mathrm{AoI}} > d\} \\
&\leq 1 - \int_0^d \Pr\{B + T^{\mathrm{I}}(k-1,k) \leq x\} f_{T^{\mathrm{S}}}(d-x)dx \\
&= 1 - (0 + \int_D^d (1 - e^{-\theta(x-D)}) \cdot \frac{1}{\mu} e^{-\frac{1}{\mu}(d-x)} dx) \\
&= e^{-\frac{1}{\mu}(d-D)} + \frac{1}{1-\theta\mu}(e^{\theta(D-d)} - e^{-\frac{1}{\mu}(d-D)})
\end{aligned} \tag{28}$$

Here, θ can be optimized according to Eq. (22), there holds

$$\mathrm{E}[e^{\theta T^{\mathrm{S}}(1)}]\mathrm{E}[e^{-\theta T^{\mathrm{I}}(0,1)}] = \frac{1}{1-\mu\theta} e^{-\theta D} \leq 1 \tag{29}$$

In this case, a closed-form optimal θ^* is unavailable while the analytical value can be obtained with the help of calculation tool like MATLAB. Hence, while $d > D$, we finally have:

$$\Pr\{P^{\mathrm{AoI}} > d\} \leq e^{-\frac{1}{\mu}(d-D)} + \frac{e^{\theta^*(D-d)} - e^{-\frac{1}{\mu}(d-D)}}{1-\theta^*\mu} \tag{30}$$

where θ^* is the maximum allowable value satisfying Eq. (29).

As a result, an upper bound of the average peak AoI of D/M/1 queueing model can be obtained as

$$\begin{aligned}
\mathrm{E}[P^{\mathrm{AoI}}] &= \int_0^D \Pr\{P^{\mathrm{AoI}} > d\}dd + \int_D^\infty \Pr\{P^{\mathrm{AoI}} > d\}dd \\
&= D + \int_D^\infty e^{-\frac{1}{\mu}(d-D)} + \frac{e^{\theta^*(D-d)} - e^{-\frac{1}{\mu}(d-D)}}{1-\theta^*\mu} dd \\
&= D + \mu + \frac{1}{\theta^*}
\end{aligned} \tag{31}$$

IV. Numerical Results

In this section, numerical results are provided and discussed for the two cases, M/M/1 and D/M/1. The former M/M/1 case has often been adopted in the literature for AoI performance analysis, e.g., see [10] and [15]. The latter case D/M/1 corresponds to a practical setting where the updates are sent periodically. The system utilization, defined as $\rho = \frac{\mathrm{E}[T^{\mathrm{S}}(k)]}{\mathrm{E}[T^{\mathrm{I}}(k-1,k)]}$ $(0 < \rho < 1)$, will also be used as a performance parameter.

Fig. 2 and Fig. 3 depict the peak AoI violation probability varying with the server utilization ρ and the maximum tolerant threshold d based on M/M/1 and D/M/1 queuing models, respectively. We set the average transmission time for any update as $\mu = 1$ (time unit), then ρ is equal to the reciprocal of average inter-arrival time (λ or D). It is observed that the peak AoI violation probability decreases as the maximum tolerant threshold increase with a given server utilization. Meanwhile, for any maximum tolerant threshold, there also exits an optimal server utilization setting to minimize the peak AoI violation probability. This is because the peak AoI depends on inter-arrival time, queueing time and transmission time of an update. With a given service capability, the statistical characteristics of the transmission time is usually fixed. In this regard, low server utilization means huge inter-arrival time. Larger inter-arrival time may result in few updates transmitted to the receiver, which deteriorates the peak AoI performance. Additionally, smaller inter-arrival time may result in high queuing delay for updates, which also has negative impact on the peak AoI performance. Therefore, it is critical to design an appropriate server utilization scheme for peak AoI guarantee.

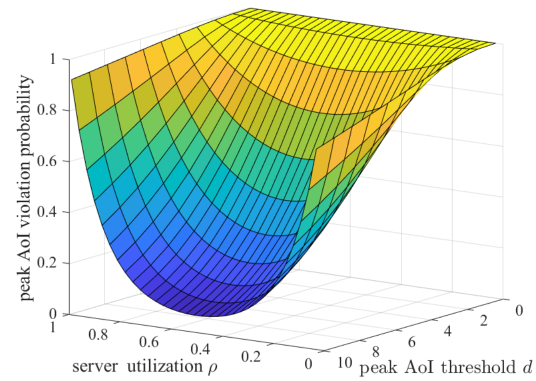

Fig. 2. Peak AoI violation probability of the M/M/1 case.

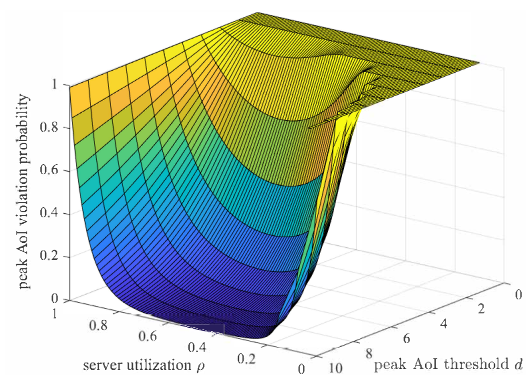

Fig. 3. Peak AoI violation probability of the D/M/1 case.

Fig. 4 depicts the relationship between peak AoI violation probability and server utilization under different service capability, where the threshold of peak AoI set as $d = 3$ (time unit). In addition to the observation from Fig. 2 and Fig. 3, it is found that the optimal server utilization of M/M/1 is irrelevant with the service capability. Specifically, the optimal utilization is also equal to 0.5. Differently, the optimal server utilization of D/M/1 queuing model is positively correlated with the average transmission time μ. Additionally, Fig. 4 verifies that the peak AoI violation probability can be reduced by decreasing the transmission time of an update.

Fig. 5 depicts the relationship between the average peak AoI and the server utilization. In addition, the obtained performance bound is compared with exact results obtained from [10] and [15] based on the classical queueing theory

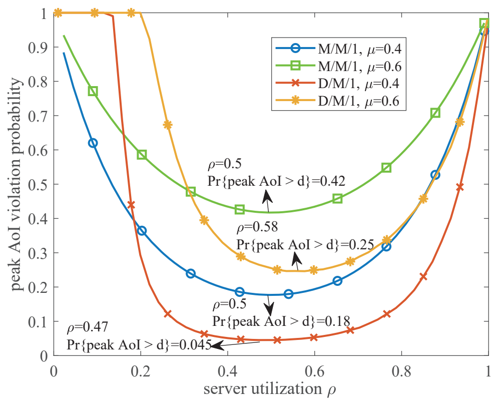

Fig. 4. The influence of server utilization on the peak AoI violation probability.

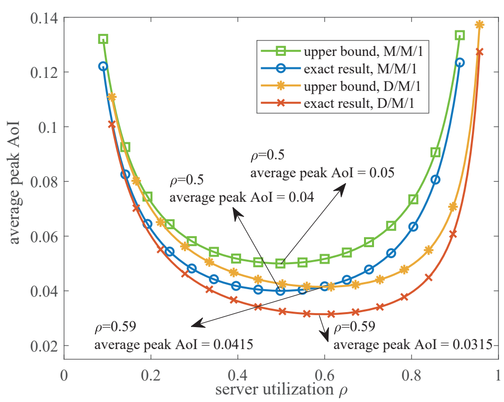

Fig. 5. The influence of server utilization on the average peak AoI.

analysis. We set the average transmission time of the system as $\mu = 0.01s$, which reflects the actual multimedia transmission. It is observed that there exists optimal configuration of server utilization in average peak AoI minimization for each case. Also, the optimal server utilization or the status update arrival obtained from the upper bound is identical to that obtained from the exact result. Additionally, it can be found that the obtained performance bounds for the M/M/1 case and D/M/1 case are closed to the corresponding exact result, respectively. Specifically, the gaps between the upper bound and exact result for both cases are always one transmission time μ, which validates the tightness of the upper bound.

V. Conclusions

In this paper, the probabilistic characteristics of the peak AoI have been studied. Different from the literature where the focus is mostly on average (peak) AoI, we provided an analysis on peak AoI distribution. By decoupling the inter-arrival time and the sojourn time of each update, the peak AoI can be expressed with the inter-arrival time and the transmission time of each update. With the help of martingale theory, an upper bound on peak AoI violation probability was derived for the general GI/GI/1 setting, and applied to two specific cases, namely M/M/1 and D/M/1. The impact of server utilization on peak AoI performance was also investigated. Additionally, numerical results verifying the validity of the proposed peak AoI bound were presented and discussed.

VI. Acknowledgement

This work was supported in part by the National Natural Science Foundation of China under grants 61901078, and in part by Chongqing Postgraduate Research and Innovation Project under grant CYB22250, and in part by the China Postdoctoral Science Foundation under grant 2022MD713692, and in part by the Chongqing Postdoctoral Science Special Foundation under grant 2021XM2018, and in part by the Natural Science Foundation of Chongqing under grant cstc2020jcyj-zdxmX0024, and in part by University Innovation Research Group of Chongqing under grant CXQT20017, and in part by the Youth Innovation Group Support Program of ICE Discipline of CQUPT under grant SCIE-QN-2022-04.

References

[1] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1807–1827, 2019.

[2] M. Noroozi and M. Fidler, "A min-plus model of age-of-information with worst-case and statistical bounds," in *ICC 2022 - IEEE International Conference on Communications*, 2022, pp. 2090–2095.

[3] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1183–1210, 2021.

[4] F. Chiariotti, O. Vikhrova, B. Soret, and P. Popovski, "Peak age of information distribution for edge computing with wireless links," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3176–3191, 2021.

[5] B. Zhou and W. Saad, "On the age of information in internet of things systems with correlated devices," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[6] J. P. Champati, R. R. Avula, T. J. Oechtering, and J. Gross, "Minimum achievable peak age of information under service preemptions and request delay," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 5, pp. 1365–1379, 2021.

[7] Y. Gu, H. Chen, C. Zhai, Y. Li, and B. Vucetic, "Minimizing age of information in cognitive radio-based IoT systems: Underlay or overlay?" *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 273–10 288, 2019.

[8] Z. Fang, J. Wang, Y. Ren, Z. Han, H. V. Poor, and L. Hanzo, "Age of information in energy harvesting aided massive multiple access networks," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 5, pp. 1441–1456, 2022.

[9] 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 Transactions on Information Theory*, vol. 65, no. 12, pp. 8305–8324, 2019.

[10] L. Hu, Z. Chen, Y. Dong, Y. Jia, L. Liang, and M. Wang, "Status update in IoT networks: Age-of-information violation probability and optimal update rate," *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11 329–11 344, 2021.

[11] N. Akar and O. Dogan, "Discrete-time queueing model of age of information with multiple information sources," *IEEE Internet of Things Journal*, vol. 8, no. 19, pp. 14 531–14 542, 2021.

[12] X. Zhang, J. Wang, and H. V. Poor, "Aoi-driven statistical delay and error-rate bounded QoS provisioning for mURLLC over UAV-multimedia 6G mobile networks using FBC," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3425–3443, 2021.

[13] Y. Jiang, "Network calculus and queueing theory: two sides of one coin," 2010.

[14] J. L. Doob, *Stochastic processes*. New York, NY, USA: Wiley, 1953.

[15] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends in Networking*, vol. 12, no. 3, pp. 162–259, 2017.