AoI Perspective on the Accuracy of Monitoring Systems for Continuous-Time Markovian Sources

Yoshiaki Inoue and Tetsuya Takine
Department of Information and Communications Technology, Osaka University, Suita 565-0871, Japan
E-mail:{yoshiaki, takine}@comm.eng.osaka-u.ac.jp

***Abstract*—We consider a remote monitoring system where a monitor displays the latest state information obtained from a time-varying information source. Assuming that the information source is represented as a continuous-time Markov chain, we analyze the effect of the age of information (AoI) on the accuracy of the monitoring system. We first obtain the conditional probability of the displayed state, given the actual current state of the information source. We then derive simple lower and upper bounds for the probability that the actual current state is displayed. Finally, we develop a computing method of these probabilities for a special case that the information source is represented as a reversible Markov chain.**

I. Introduction

We consider a situation that the state of a time-varying information source is monitored remotely. Specifically, the information source is attached with a sensor node, which sends observed data to a remote server. The server processes the received data and sends extracted state information to a monitor that displays the latest state information received. Because the state of the information source changes over time, the value of displayed information degenerates with time. It is thus important for such a system that the displayed information on the monitor is kept sufficiently fresh.

The age of information (AoI) [1] is a widely used metric of the freshness of information, which is defined as the elapsed time from the generation time of currently displayed information. Specifically, the AoI A_t at time t is defined as

$$A_t = t - \eta_t, \quad t \in \mathbb{R}, \tag{1}$$

where η_t denotes the generation time of the displayed information at time t. The mean AoI $\mathrm{E}[A]$ is the primary performance measure of interest, which is defined as

$$\mathrm{E}[A] = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} A_t \mathrm{d}t.$$

Under a fairly general setting, the mean AoI $\mathrm{E}[A]$ is given in terms of the mean inter-sampling time, the mean delay, and a cross-term of the inter-sampling time and the delay [1].

To characterize the delay formally, the system is usually modeled as a queueing system. Based on the general formula, various analytical results for the mean AoI $\mathrm{E}[A]$ has been derived in the literature. In [1], first-come first-served (FCFS) queues are analyzed. Last-come first-served (LCFS) queues are considered in [2] and [3], where newer data is given priority in transmission and processing. A multi-class FCFS queue is considered in [4], which is extended to a system with priority mechanism in [5]. To model a situation that the sensor node communicates with the server through a network, multi-server queueing models are considered in [6].

In designing a monitoring system, one needs to keep the AoI *sufficiently small*, so that the monitor accurately displays the current state of the information source. Intuitively, a fairly large value of the AoI would be acceptable if the information source is slowly varying in time, and otherwise, a strict limit for the AoI would be imposed. Therefore, a target value of the AoI (or precisely, target statistical properties of the AoI such as the mean and variance) is highly dependent on the nature of *the information source dynamics*.

There are only a few works that discuss relations between the AoI and the dynamics of the information source in the literature. The age of channel state information of a wireless link is considered in [7], where the channel state is modeled as a discrete-time two-state Markov chain. In [8], the information source is assumed to be a Wiener process, and an optimal sampling policy minimizing the mean square error between the actual and displayed states is derived. It is also shown in [8] that if the AoI process is independent of the monitored Wiener process, the mean squared error is equal to the mean AoI $\mathrm{E}[A]$. In [9], the mutual information between the actual and displayed states is proposed as a metric for the freshness of information, assuming a discrete-time Markov chain for the information source. Solving an optimal sampling problem under a more general setting, a sampling policy maximizing the mutual information is derived in [9].

In this paper, we consider a continuous-time monitoring system where the information source is represented as a finite-state Markov chain, and the AoI process $(A_t)_{t\in\mathbb{R}}$ is an ergodic stochastic process independent of the information source. As we will see, the information aging in our system is characterized by a matrix (denoted by $\boldsymbol{R}$), whose (i, j)th element represents the probability that state j is displayed on the monitor, given that the actual current state is equal to i. The main contributions of this paper are summarized as follows:

(i) We derive an expression for the matrix $\boldsymbol{R}$, in terms of the *probability distribution* of the AoI [10].

(ii) We derive lower and upper bounds for diagonal elements of $\boldsymbol{R}$, i.e., the conditional probabilities that the monitor correctly displays the current state of the information source, given the actual current state. In particular, the lower bound is given in terms of the *product of the mean AoI and information source's transition rate*, which provides a simple criteria to ensure the performance of the monitoring system.

(iii) For a special case that the Markov chain is reversible, we develop a computing method for the matrix $\boldsymbol{R}$. We also provide some numerical examples, and discuss the gap between the lower bound and diagonal elements of $\boldsymbol{R}$.

The rest of this paper is organized as follows. In Section II, we formally describe our system model. In Section III, we present analytical results for a general case, and develop a computing method for the reversible case. In Section IV, we provide some numerical examples. Finally, we conclude this paper in Section V.

II. Model

A. Information source

We consider an information source whose state takes a value in a finite set $\mathcal{M} = \{1, 2, \dots, M\}$. Let Y_t ($t \in \mathbb{R}$) denote the state of the information source at time t. Throughout this paper, we assume the following:

Assumption 1. *$(Y_t)_{t\in\mathbb{R}}$ is stationary and ergodic.*

We call the information source satisfying Assumption 1 *the general information source*. Let $\boldsymbol{\pi} := (\pi_1, \pi_2, \dots, \pi_M)$ denote the stationary probability vector of $(Y_t)_{t\in\mathbb{R}}$:

$$\pi_i = \Pr(Y_t = i), \quad i \in \mathcal{M}.$$

We also consider *a Markovian information source*, where $(Y_t)_{t\in\mathbb{R}}$ forms a finite-state continuous-time Markov chain [11, Chapter 8] satisfying Assumption 1. Note that the Markov chain $(Y_t)_{t\in\mathbb{R}}$ is characterized by transition rates $q_{i,j}$ ($i \in \mathcal{M}$, $j \in \mathcal{M}$, $j \neq i$), where $q_{i,j} \geq 0$. We define q_i ($i \in \mathcal{M}$) as

$$q_i = \sum_{j\in\mathcal{M}, j\neq i} q_{i,j}.$$

The Markov chain stays at state i ($i \in \mathcal{M}$) for exponentially distributed length of time with mean $1/q_i$, and then transitions to state j ($j \in \mathcal{M}$, $j \neq i$) with probability $q_{i,j}/q_i$.

Let $\boldsymbol{Q}$ denote the infinitesimal generator of $(Y_t)_{t\in\mathbb{R}}$, which is an $M \times M$ matrix whose (i, j)th element is given by

$$[\boldsymbol{Q}]_{i,j} = \begin{cases} -q_i & j = i, \\ q_{i,j} & j \neq i. \end{cases}$$

By definition, $\boldsymbol{Q}$ has negative diagonal elements and nonnegative offdiagonal elements, which satisfy

$$\boldsymbol{Q}\boldsymbol{e} = \boldsymbol{0},$$

where $\boldsymbol{e} := (1, 1, \dots, 1)^\top$ denotes an $M \times 1$ vector of ones.

It is well known that the transition probabilities of this Markov chain is given by a matrix exponential function:

$$\Pr(Y_{t+x} = j \mid Y_t = i) = \big[\exp[\boldsymbol{Q}x]\big]_{i,j}, \quad t \in \mathbb{R}, x \geq 0.$$

Note that the transition probabilities do not depend on the initial time t. We define $\boldsymbol{P}(x)$ ($x \geq 0$) as the transition probability matrix of this Markov chain:

$$\boldsymbol{P}(x) = \exp[\boldsymbol{Q}x]. \tag{2}$$

Note that the stationary probability vector $\boldsymbol{\pi}$ is uniquely determined by the balance equation $\boldsymbol{\pi}\boldsymbol{Q} = \boldsymbol{0}$, $\boldsymbol{\pi}\boldsymbol{e} = 1$.

B. Monitoring system

The state $(Y_t)_{t\in\mathbb{R}}$ of the information source is sampled at time $\hat{t}_n$ ($n \in \mathbb{Z}$) and it is notified to the monitor at t_n, where $\hat{t}_n < \hat{t}_{n+1}$, $t_n < t_{n+1}$, and $\hat{t}_n \leq t_n$ for all $n \in \mathbb{Z}$. Without loss of generality, we assume $t_0 = 0$. Let $\eta_t := \hat{t}_{\sup\{n;\, t_n \leq t\}}$ ($t \in \mathbb{R}$) denote the generation time of the information displayed on the monitor at time t. The AoI A_t ($t \in \mathbb{R}$) at time t is then given by (1).

Let $\hat{Y}_t$ ($t \in \mathbb{R}$) denote the state of the information source displayed on the monitor at time t:

$$\hat{Y}_t = Y_{t-A_t}.$$

We make the following assumptions.

Assumption 2. *$(A_t)_{t\in\mathbb{R}}$ is stationary and ergodic.*

Assumption 3. *$(Y_t)_{t\in\mathbb{R}}$ and $(A_t)_{t\in\mathbb{R}}$ are independent and jointly ergodic [12].*

Remark 1. *$(Y_t)_{t\in\mathbb{R}}$ and $(A_t)_{t\in\mathbb{R}}$ are jointly ergodic if $(Y_t)_{t\in\mathbb{R}}$ is mixing and $(A_t)_{t\in\mathbb{R}}$ is ergodic [12, Theorem 2]. Note that the ergodic Markovian information source is mixing.*

Remark 2. *Under Assumptions 1 to 3, the following relation holds sample-path wise with probability one [13, Page 50]:*

$$\mathrm{E}[\phi(A_0, Y_0)] = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \phi(A_t, Y_t)\mathrm{d}t,$$

where $\phi(x, y)$ ($x \geq 0$, $y \geq 0$) denotes any non-negative function for which $\mathrm{E}[\phi(A_0, Y_0)]$ exists.

Because of the stationarity, the distribution of the AoI A_t ($t \in \mathbb{R}$) is independent of t. Let A denote a generic random variable following the probability distribution of A_t. We define $A(x) := \Pr(A \leq x)$ ($x \geq 0$) as the probability distribution function of A.

C. Performance measure

To evaluate the accuracy of the monitoring system, we consider the conditional probability that state j is displayed on the monitor, given that the actual current state of the information source is equal to i:

$$r_{i,j} := \Pr(\hat{Y}_t = j \mid Y_t = i), \quad i \in \mathcal{M}, j \in \mathcal{M}.$$

From Remark 2, we can verify that the following relation holds sample-path wise with probability one:

$$r_{i,j} = \lim_{T\to\infty} \frac{\displaystyle\int_{-T/2}^{T/2} \mathbb{1}_{\{Y_t=i\}} \mathbb{1}_{\{\hat{Y}_t=j\}} \mathrm{d}t}{\displaystyle\int_{-T/2}^{T/2} \mathbb{1}_{\{Y_t=i\}} \mathrm{d}t}. \tag{3}$$

$r_{i,j}$ thus represents the long-run fraction of time that state j is displayed, given that the actual state is equal to i.

Various performance measures for the monitoring system can be represented in terms of $r_{i,j}$ ($i \in \mathcal{M}$, $j \in \mathcal{M}$).

Example 1. *In [9], the mutual information of $(Y_t)_{t\in\mathbb{R}}$ and $(\hat{Y}_t)_{t\in\mathbb{R}}$ is proposed as a metric for the freshness of information. By definition, the mutual information is given by*

$$\sum_{i\in\mathcal{M}}\sum_{j\in\mathcal{M}} \pi_i r_{i,j} \log\left(\frac{\pi_i r_{i,j}}{\pi_i \hat{\pi}_j}\right),$$

where $\hat{\pi}_j := \Pr(\hat{Y} = j)$. We will show $\hat{\pi}_j = \pi_j$ in Lemma 1.

Example 2. *Assume that we have a distance function $d(i,j)$ ($i \in \mathcal{M}$, $j \in \mathcal{M}$) for the state space $\mathcal{M}$. The following performance measures quantify the accuracy of the monitoring system.*

- *Mean error between $(Y_t)_{t\in\mathbb{R}}$ and $(\hat{Y}_t)_{t\in\mathbb{R}}$:*

$$\sum_{i\in\mathcal{M}}\sum_{j\in\mathcal{M}} \pi_i r_{i,j} d(i,j).$$

- *Maximum mean error between $(Y_t)_{t\in\mathbb{R}}$ and $(\hat{Y}_t)_{t\in\mathbb{R}}$:*

$$\max_{i\in\mathcal{M}} \sum_{j\in\mathcal{M}} r_{i,j} d(i,j).$$

- *Probability that the error between $(Y_t)_{t\in\mathbb{R}}$ and $(\hat{Y}_t)_{t\in\mathbb{R}}$ exceeds a given threshold d_{th}:*

$$\sum_{i\in\mathcal{M}} \sum_{j\in\{k\in\mathcal{M};\, d(i,k)>d_{\mathrm{th}}\}} \pi_i r_{i,j}.$$

Associated with $r_{i,j}$, we also define $\overline{r}_{i,j}$ as the probability that the current state of the information source is equal to j, given that state i is displayed on the monitor:

$$\overline{r}_{i,j} = \Pr(Y_t = j \mid \hat{Y}_t = i), \quad i \in \mathcal{M}, j \in \mathcal{M}.$$

Similarly to (3), we have

$$\overline{r}_{i,j} = \lim_{T\to\infty} \frac{\displaystyle\int_{-T/2}^{T/2} \mathbb{1}_{\{Y_t=j\}} \mathbb{1}_{\{\hat{Y}_t=i\}} \mathrm{d}t}{\displaystyle\int_{-T/2}^{T/2} \mathbb{1}_{\{\hat{Y}_t=i\}} \mathrm{d}t},$$

i.e., $\overline{r}_{i,j}$ is equal to the long-run fraction of time that the actual current state is equal to j, given that state i is displayed.

While $r_{i,j}$ is considered to be a performance measure representing the *accuracy* of the monitoring system for state i, $\overline{r}_{i,j}$ provides an *estimate* for the current state given the displayed state i. We define $\boldsymbol{R}$ and $\overline{\boldsymbol{R}}$ as $M \times M$ matrices whose (i,j)th elements are given by $r_{i,j}$ and $\overline{r}_{i,j}$:

$$[\boldsymbol{R}]_{i,j} = r_{i,j}, \quad [\overline{\boldsymbol{R}}]_{i,j} = \overline{r}_{i,j}, \quad i \in \mathcal{M}, j \in \mathcal{M}.$$

III. Main Results

A. A monitoring system for a general information source

We first consider a monitoring system for a general information source.

Lemma 1. *Under Assumptions 1 to 3, $(\hat{Y}_t)_{t\in\mathbb{R}}$ is a stationary stochastic process with*

$$\Pr(\hat{Y}_t = i) = \pi_i, \quad i \in \mathcal{M}. \tag{4}$$

Proof. For $x_1 \le x_2 \cdots \le x_n$ and $i_\ell \in \mathcal{M}$ ($\ell = 1, 2, \dots, n$),

$$\begin{aligned}
&\Pr(\hat{Y}_{t+x_\ell} = i_\ell, \ell \in \{1,2,\dots,n\}) \\
&\quad = \Pr(Y_{t+x_\ell - A_{t+x_\ell}} = i_\ell, \ell \in \{1,2,\dots,n\}) \\
&\quad = \Pr(Y_{x_\ell - A_{x_\ell}} = i_\ell, \ell \in \{1,2,\dots,n\}) \\
&\quad = \Pr(\hat{Y}_{x_\ell} = i_\ell, \ell \in \{1,2,\dots,n\}),
\end{aligned}$$

where the second equality follows from the assumption that $(Y_t)_{t\in\mathbb{R}}$ and $(A_t)_{t\in\mathbb{R}}$ are independent stationary stochastic processes. We also obtain (4) from the stationarity and independence of $(Y_t)_{t\in\mathbb{R}}$ and $(A_t)_{t\in\mathbb{R}}$:

$$\begin{aligned}
\Pr(\hat{Y}_t = i) &= \int_0^\infty \Pr(Y_{t-x} = i) \mathrm{d}A(x) \\
&= \Pr(Y_t = i) \int_0^\infty \mathrm{d}A(x) \\
&= \pi_i.
\end{aligned}$$

□

Remark 3. *From Remark 2 and Lemma 1, the following relation holds sample-path wise with probability one:*

$$\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathbb{1}_{\{\hat{Y}_t=i\}} \mathrm{d}t = \Pr(\hat{Y}_0 = i) = \pi_i, \quad i \in \mathcal{M}.$$

Theorem 1. *Under Assumptions 1 to 3, $r_{i,j}$ and $\overline{r}_{i,j}$ ($i \in \mathcal{M}$, $j \in \mathcal{M}$) are related by*

$$r_{i,j} = \frac{\pi_j \overline{r}_{j,i}}{\pi_i}, \quad i \in \mathcal{M}, j \in \mathcal{M}. \tag{5}$$

In particular, we have

$$r_{i,i} = \overline{r}_{i,i}, \quad i \in \mathcal{M}. \tag{6}$$

Proof. (5) follows from Lemma 1 and Bayes' formula, and (6) follows from (5). □

In view of (6), we introduce $\{r_i\}_{i\in\mathcal{M}}$ for $r_{i,i}$ and $\overline{r}_{i,i}$.

$$r_i := r_{i,i} = \overline{r}_{i,i}, \quad i \in \mathcal{M}. \tag{7}$$

Note that r_i ($i \in \mathcal{M}$) is our primary quantity of interest.

B. A monitoring system for a Markovian information source

In this subsection, we consider a monitoring system for a Markovian information source under Assumptions 1 to 3. We rewrite (5) in a matrix form:

$$\boldsymbol{R} = \boldsymbol{\Pi}^{-1} \overline{\boldsymbol{R}}^\top \boldsymbol{\Pi}, \tag{8}$$

where $\boldsymbol{\Pi}$ denotes an $M \times M$ diagonal matrix whose ith ($i \in \mathcal{M}$) diagonal element is given by π_i.

Theorem 2. *$\boldsymbol{R}$ and $\overline{\boldsymbol{R}}$ are given by*

$$\boldsymbol{R} = \int_0^\infty \exp\left[(\boldsymbol{\Pi}^{-1}\boldsymbol{Q}^\top\boldsymbol{\Pi})x\right]\mathrm{d}A(x), \tag{9}$$

$$\overline{\boldsymbol{R}} = \int_0^\infty \exp[\boldsymbol{Q}x]\mathrm{d}A(x). \tag{10}$$

Remark 4. *$\boldsymbol{\Pi}^{-1}\boldsymbol{Q}^\top\boldsymbol{\Pi}$ in (9) can be regarded as the infinitesimal generator of the time-reversed process $(Y_{-t})_{t\in\mathbb{R}}$ of the Markovian information source [14, Page 28].*

Proof. Noting (8), we can verify that (10) implies (9). It thus suffices to prove (10). From Assumptions 2 and 3, we have

$$\begin{aligned}\overline{r}_{i,j} &= \Pr(Y_t = j \mid Y_{t-A_t} = i)\\ &= \int_{x=0}^\infty \Pr(Y_t = j \mid Y_{t-x} = i)\mathrm{d}A(x)\\ &= \int_{x=0}^\infty [\boldsymbol{P}(x)]_{i,j}\mathrm{d}A(x).\end{aligned} \tag{11}$$

(10) thus follows from (2). □

As stated in Theorem 1, $\boldsymbol{R}$ and $\overline{\boldsymbol{R}}$ have the same diagonal elements $\{r_i\}_{i\in\mathcal{M}}$, which can also be verified by Theorem 2. Below we derive lower and upper bounds for r_i ($i \in \mathcal{M}$).

Let $a^*(s)$ ($\mathrm{Re}(s) > 0$) denote the Laplace-Stieltjes transform (LST) of the AoI distribution:

$$a^*(s) = \mathrm{E}[\exp[-sA]] = \int_0^\infty \exp[-sx]\mathrm{d}A(x).$$

Theorem 3. *r_i ($i \in \mathcal{M}$) in (7) is bounded by*

$$a^*(q_i) \le r_i \le 1 - (a^*(q_{\max}) - a^*(q_i + q_{\max})), \tag{12}$$

where $q_{\max} := \max_{i\in\mathcal{M}} q_i$.

Proof. We first consider the lower bound. By definition,

$$\begin{aligned}[\boldsymbol{P}(x)]_{i,i} &= \Pr(Y_x = i \mid Y_0 = i)\\ &\ge \Pr(Y_t = i \text{ for } t \in [0,x] \mid Y_0 = i) = \exp[-q_i x].\end{aligned}$$

We then obtain the lower bound in (12) from (7) and (11).

We next consider the upper bound. We rewrite $[\boldsymbol{P}(x)]_{i,i}$ as

$$\begin{aligned}[\boldsymbol{P}(x)]_{i,i} &= 1 - \sum_{j\in\mathcal{M},j\neq i}[\exp[\boldsymbol{Q}x]]_{i,j}\\ &= 1 - \sum_{j\in\mathcal{M},j\neq i}\int_0^x q_i\exp[-q_i t]\\ &\qquad \cdot \sum_{k\in\mathcal{M},k\neq i}\frac{q_{i,k}}{q_i}[\boldsymbol{P}(x-t)]_{k,j}\mathrm{d}t.\end{aligned}$$

From this equation, we have

$$\begin{aligned}&[\boldsymbol{P}(x)]_{i,i}\\ &\le 1 - \sum_{k\in\mathcal{M},k\neq i}\int_0^x q_i\exp[-q_i t]\cdot\frac{q_{i,k}}{q_i}\cdot[\boldsymbol{P}(x-t)]_{k,k}\mathrm{d}t\\ &\le 1 - \sum_{k\in\mathcal{M},k\neq i}\frac{q_{i,k}}{q_i}\int_0^x q_i\exp[-q_i t]\exp[-q_k(x-t)]\mathrm{d}t\\ &\le 1 - \sum_{k\in\mathcal{M},k\neq i}\frac{q_{i,k}}{q_i}\int_0^x q_i\exp[-q_i t]\exp[-q_k x]\mathrm{d}t\\ &= 1 - \sum_{k\in\mathcal{M},k\neq i}\exp[-q_k x]\cdot\frac{q_{i,k}}{q_i}(1-\exp[-q_i x])\\ &\le 1 - \exp[-q_{\max}x](1-\exp[-q_i x])\sum_{k\in\mathcal{M},k\neq i}\frac{q_{i,k}}{q_i}\\ &= 1 - (\exp[-q_{\max}x] - \exp[-(q_i+q_{\max})x]).\end{aligned}$$

We thus obtain the upper bound in (12) from (7) and (11). □

Corollary 1. *If $\mathrm{E}[A] < \infty$, r_i is bounded by*

$$1 - q_i\mathrm{E}[A] \le r_i \le 1 - q_i\mathrm{E}[A] + \frac{(q_i+q_{\max})^2}{2}\mathrm{E}[A^2]. \tag{13}$$

Remark 5. *The lower bound in (13) offers a simple criteria to ensure the accuracy of the monitoring system. Specifically, in order to ensure $r_i \ge 1-\epsilon$ for some $\epsilon > 0$, it is sufficient to design the system so that $\mathrm{E}[A] \le \epsilon/q_i$.*

Proof. Note that for any $\theta \ge 0$ and $x \ge 0$,

$$1 - \theta x \le \exp[-\theta x] \le 1 - \theta x + \frac{\theta^2 x^2}{2},$$

so that

$$1 - \theta\mathrm{E}[A] \le a^*(\theta) \le 1 - \theta\mathrm{E}[A] + \frac{\theta^2\mathrm{E}[A^2]}{2}.$$

Using this relation, we can derive (13) from (12) with straightforward calculations. □

C. A special case: reversible Markovian information source

In this subsection, we develop a computing method for $\boldsymbol{R}$, assuming that the Markovian information source $(Y_t)_{t\in\mathbb{R}}$ is *reversible* [14], i.e., its time-reversed process $(Y_{-t})_{t\in\mathbb{R}}$ follows the same probability law as the original process $(Y_t)_{t\in\mathbb{R}}$. Specifically, we make the following assumption in addition to Assumptions 1 to 3.

Assumption 4. *In the Markov chain $(Y_t)_{t\in\mathbb{R}}$, detailed balance equations hold:*

$$\pi_i q_{i,j} = \pi_j q_{j,i}, \quad i \in \mathcal{M}, j \in \mathcal{M}. \tag{14}$$

It is known that (14) is a necessary and sufficient condition for the Markov chain being reversible [14]. Note that we can rewrite (14) as $\boldsymbol{\Pi}\boldsymbol{Q} = \boldsymbol{Q}^\top\boldsymbol{\Pi}$, which implies $\boldsymbol{Q} = \boldsymbol{\Pi}^{-1}\boldsymbol{Q}^\top\boldsymbol{\Pi}$ (cf. Remark 4). Therefore, it follows from (9) and (10) that

$$\boldsymbol{R} = \overline{\boldsymbol{R}}, \tag{15}$$

which is almost obvious from the time-reversibility.

Let $\boldsymbol{D}$ denote an $M\times M$ diagonal matrix whose ith ($i \in \mathcal{M}$) element is given by $\sqrt{\pi_i}$. We define an $M \times M$ matrix $\boldsymbol{S}$ as

$$\boldsymbol{S} = \boldsymbol{D}\boldsymbol{Q}\boldsymbol{D}^{-1}. \tag{16}$$

It follows from (14) that $\boldsymbol{S}$ is a real symmetric matrix, which can be verified with $\boldsymbol{Q}^\top = \boldsymbol{\Pi}\boldsymbol{Q}\boldsymbol{\Pi}^{-1}$:

$$\boldsymbol{S}^\top = \boldsymbol{D}^{-1}\boldsymbol{Q}^\top\boldsymbol{D} = \boldsymbol{D}^{-1}(\boldsymbol{\Pi}\boldsymbol{Q}\boldsymbol{\Pi}^{-1})\boldsymbol{D} = \boldsymbol{D}\boldsymbol{Q}\boldsymbol{D}^{-1} = \boldsymbol{S}.$$

Owing to the properties of real symmetric matrices, all eigenvalues of $\boldsymbol{S}$ are real-valued, and $\boldsymbol{S}$ is diagonizable by an orthogonal matrix $\boldsymbol{U} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M)$, where $\boldsymbol{u}_k$ ($k \in \mathcal{M}$)

denotes a normalized right eigenvector (i.e., $\boldsymbol{u}_k^\top \boldsymbol{u}_k = 1$) of $\boldsymbol{S}$ associated with the kth largest eigenvalue γ_k:

$$\boldsymbol{S} = \sum_{k=1}^{M} \gamma_k \boldsymbol{u}_k \boldsymbol{u}_k^\top.$$

Note that $\boldsymbol{Q}$ and $\boldsymbol{S}$ have the same set of eigenvalues. It then follows from Perron-Frobenius theorem that $\gamma_1 = 0$ and $\gamma_k < 0$ $(k = 2, 3, \dots, M)$. We then define $\theta_k := -\gamma_k$. By definition, $\theta_k > 0$ $(k = 2, 3, \dots, M)$, and $\boldsymbol{Q}$ is rewritten as

$$\boldsymbol{Q} = \sum_{k=2}^{M} (-\theta_k) \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D}. \tag{17}$$

Remark 6. *$\boldsymbol{D}^{-1}\boldsymbol{u}_k$ (resp. $\boldsymbol{u}_k^\top \boldsymbol{D}$) denotes a right (resp. left) eigenvector of $\boldsymbol{Q}$ associated with the kth largest eigenvalue $-\theta_k$. In particular, we can verify that for some constant $c > 0$,*

$$\boldsymbol{D}^{-1}\boldsymbol{u}_1 = c\boldsymbol{e}, \qquad \boldsymbol{u}_1^\top \boldsymbol{D} = (1/c)\boldsymbol{\pi}. \tag{18}$$

Theorem 4. *If $(Y_t)_{t\in\mathbb{R}}$ is a reversible Markov chain, $\boldsymbol{R}$ is given in terms of the LST $a^*(s)$ of the AoI distribution by*

$$\boldsymbol{R} = \boldsymbol{e}\boldsymbol{\pi} + \sum_{k=2}^{M} a^*(\theta_k) \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D}. \tag{19}$$

Proof. It follows from (17) that

$$\begin{aligned}
\exp[\boldsymbol{Q}x] &= \boldsymbol{I} + \sum_{n=1}^{\infty} \frac{(\boldsymbol{Q}x)^n}{n!} \\
&= \boldsymbol{I} + \sum_{n=1}^{\infty} \sum_{k=2}^{M} \frac{(-\theta_k x)^n}{n!} \cdot \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D} \\
&= \boldsymbol{I} + \sum_{k=2}^{M} (\exp[-\theta_k x] - 1) \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D} \\
&= \boldsymbol{e}\boldsymbol{\pi} + \sum_{k=2}^{M} \exp[-\theta_k x] \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D}, \qquad (20)
\end{aligned}$$

where the last equality follows from (18) and

$$\sum_{k=1}^{M} \boldsymbol{D}^{-1} \boldsymbol{u}_k \boldsymbol{u}_k^\top \boldsymbol{D} = \boldsymbol{D}^{-1} \boldsymbol{U} \boldsymbol{U}^\top \boldsymbol{D} = \boldsymbol{I}.$$

Therefore, we obtain (19) from (10) and (15). □

We summarize the computing procedure for $\boldsymbol{R}$ in Fig. 1. In step (c), we need to compute eigenvalues and eigenvectors of the real symmetric matrix $\boldsymbol{S}$, which can be easily done by means of numerical libraries (see e.g., [15]).

Remark 7. *From (20), we have*

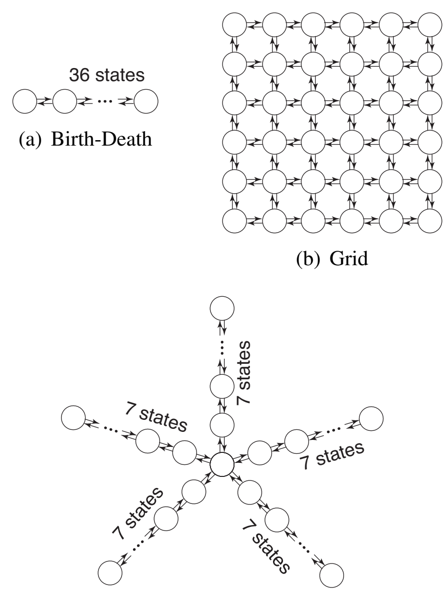
$$[\exp[\boldsymbol{Q}x]]_{i,i} = \pi_i + \sum_{k=2}^{M} b_{i,k} \exp[-\theta_k x],$$

where

$$b_{i,k} := [\boldsymbol{u}_k \boldsymbol{u}_k^\top]_{i,i} = ([\boldsymbol{u}_k]_i)^2, \quad i \in \mathcal{M}, k \in \mathcal{M}.$$

Input: $\boldsymbol{Q}$, $a^*(\theta)$ $(\theta > 0)$.
Output: $\boldsymbol{R}$.
(a) Compute the stationary probability vector $\boldsymbol{\pi}$.
(b) Compute the real symmetric matrix $\boldsymbol{S}$ from (16).
(c) Compute the eigenvalues $\{\gamma_k\}_{k\in\mathcal{M}}$ and normalized eigenvectors $\{\boldsymbol{u}_k\}_{k\in\mathcal{M}}$ of $\boldsymbol{S}$.
(d) Let $\theta_k := -\gamma_k$ $(k \in \mathcal{M})$.
(e) Compute $\boldsymbol{R}$ by (19).

Fig. 1. Computing procedure for $\boldsymbol{R}$ in the reversible case.



(a) Birth-Death (b) Grid (c) Tree

Fig. 2. Markov chains employed in numerical examples.

Because $b_{i,k} \geq 0$ $(i \in \mathcal{M}, k \in \mathcal{M})$, $[\exp[\boldsymbol{Q}x]]_{i,i}$ is a decreasing function of x, which converges to the stationary probability π_i in the limit $x \to \infty$.

In addition, it follows from (7) and (19) that

$$r_i = \pi_i + \sum_{k=2}^{M} b_{i,k} a^*(\theta_k).$$

We can verify that $b_{i,k}$ $(i \in \mathcal{M}, k \in \mathcal{M})$ satisfies

$$\begin{aligned}
\sum_{k\in\mathcal{M}} b_{i,k} &= \left[\sum_{k\in\mathcal{M}} \boldsymbol{u}_k \boldsymbol{u}_k^\top \right]_{i,i} = [\boldsymbol{U}\boldsymbol{U}^\top]_{i,i} = 1, \\
\sum_{i\in\mathcal{M}} b_{i,k} &= \sum_{i\in\mathcal{M}} ([\boldsymbol{u}_k]_i)^2 = 1,
\end{aligned}$$

i.e., $\{b_{i,k}\}_{i\in\mathcal{M},k\in\mathcal{M}}$ is doubly stochastic.

IV. Numerical Examples

In this section, we present some numerical examples. For the information source, we employ three different Markov chains with the same number of states ($M = 36$), whose transition diagrams are shown in Fig. 2. These Markov chains

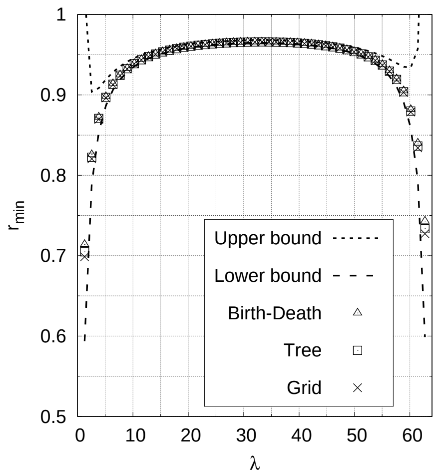


Fig. 3. $r_{\min}$ for $q = 1$ and $\mu = 64$. The bounds are computed by (13).

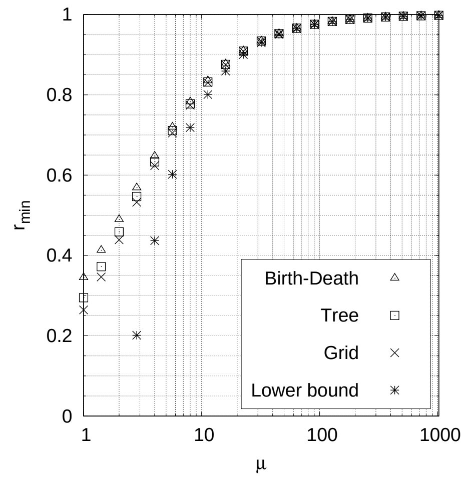


Fig. 4. $r_{\min}$ for optimal λ ($q = 1$). The Lower bound is computed by (13).

have fixed transition rate $q_i = q$ ($i \in \mathcal{M}$) and homogeneous transition probabilities $q_{i,j}/q_i = 1/\sum_{j\in\mathcal{M}} \mathbb{1}_{\{q_{i,j}\neq 0\}}$. We can verify that Assumption 4 holds for these Markov chains.

In addition, we assume that the monitoring system is formulated as a stationary FCFS D/M/1 queue, i.e., inter-sampling times are constant equal to $1/\lambda$ ($\lambda > 0$), and service times are exponentially distributed with mean $1/\mu$ ($\mu > 0$). Analytical results for the AoI distribution in the FCFS D/M/1 queue can be found in [10]:

$$a^*(s) = \left[\rho \cdot \frac{\mu - \mu\beta}{s + \mu - \mu\beta} + \tilde{g}^*(s) - \tilde{g}^*(s + \mu - \mu\beta)\right] \frac{\mu}{s+\mu},$$

$$\mathrm{E}[A] = \left(\frac{1}{2\rho} + \frac{1}{1-\beta}\right)\frac{1}{\mu},$$

$$\mathrm{E}[A^2] = \left(2\left(\frac{1}{1-\beta}\right)^2 + \frac{1}{(1-\beta)\rho} + \frac{1}{3\rho^2}\right)\left(\frac{1}{\mu}\right)^2,$$

where $\tilde{g}(s) := (1 - e^{-s/\lambda})/(s/\lambda)$, and β denotes the unique solution of $x = g^*(\mu - \mu x)$ ($0 < x < 1$).

Let $r_{\min} := \min_{i\in\mathcal{M}} r_i$. In Fig. 3, $r_{\min}$ and bounds in (13) are plotted as functions of λ for $q = 1$ and $\mu = 64$. In this case, there is little difference in $r_{\min}$ among the three Markov chains. Also, the lower bound $1 - q\mathrm{E}[A]$ well approximates $r_{\min}$, except for extremely small or large values of λ.

For a given μ, we can numerically obtain an optimal value of λ which maximizes $r_{\min}$. We can also compute an optimal λ which minimizes the mean AoI $\mathrm{E}[A]$, or equivalently, maximizes the lower bound $1 - q\mathrm{E}[A]$. In Fig. 4, $r_{\min}$ for optimal λ and $1 - q\mathrm{E}[A]$ are plotted as functions of μ for $q = 1$. From this figure, we observe that when μ takes a small value, $r_{\min}$ takes diverse values depending on the transition structure of the information source. Also, the lower bound $1-q\mathrm{E}[A]$ greatly underestimates $r_{\min}$ when μ is small. When μ takes a large value, on the other hand, the value of $r_{\min}$ is almost independent of the transition structure, and it is well approximated by the lower bound $1 - q\mathrm{E}[A]$.

V. Conclusion

We considered the effect of the AoI on the accuracy of a monitoring system. For continuous-time Markovian information source, we derived an expression for the conditional probability of the displayed state, given the actual current state (Theorem 2). We then derived simple upper and lower bounds for the probability that the monitor displays the correct state (Corollary 1). We further developed a computing method for these probabilities in the special case that the information source is represented as a reversible Markov chain. Finally, we presented numerical examples, which suggests that the lower bound derived in this paper offers a simple yet effective criteria to ensure the accuracy of the monitoring system.

Acknowledgements

This research was supported in part by JSPS KAKENHI Grant Numbers 18K18007 and 16K00034.

References

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

[2] S. Kaul, R. Yates, and M. Gruteser, “Status updates through queues,” in *Proc. of CISS 2012*, Mar. 2012..

[3] 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.

[4] R. Yates and S. Kaul, “Real-time status updating: Multiple sources,” in *Proc. of IEEE ISIT 2012*, Jul. 2012, pp. 2666–2670.

[5] S. Kaul and R. Yates, “Age of Information: Updates with priority,” in *Proc. of IEEE ISIT 2018*, Jun. 2017, pp. 2644–2648.

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

[7] M. Costa, S. Valentin, and A. Ephremides, “On the age of channel state information for a finite-state Markov model,” in *Proc. of IEEE ICC 2015*, Jun. 2015, pp. 4101–4106.

[8] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, “Remote estimation of the Wiener process over a channel with random delay,” in *Proc. of IEEE ISIT 2017*, Jun. 2017, pp. 321–325.

[9] Y. Sun and B. Cyr, “Information aging through queues: A mutual information perspective,” in *Proc. of SPAWC 2018*, Jun. 2018.

[10] 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," available online: arxiv.org/abs/1804.06139, 2018.

[11] P. Brémaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*, Springer, New York, 1999.

[12] F. Baccelli, S. Machiraju, D. Veitch, and J. Bolot, "The role of PASTA in network measurement," *IEEE/ACM Trans. Netw.*, vol. 17, no. 4, pp. 1340–1353, Aug. 2009.

[13] F. Baccelli and P. Brémaud, *Elements of Queueing Theory, 2nd ed.*, Springer, Berlin, 2003.

[14] F. P. Kelly, *Reversibility and Stochastic Networks*, John Wiley & Sons, Chichester, 1979.

[15] M. Galassi et al., *GNU Scientific Library Reference Manual (3rd Ed.)*, http://www.gnu.org/software/gsl/.