

Modeling AoII in Push- and Pull-Based Sampling of Continuous Time Markov Chains

Ismail Cosandal¹, Nail Akar², and Sennur Ulukus¹

¹University of Maryland, College Park, MD, USA

²Bilkent University, Ankara, Türkiye

Abstract—Age of incorrect information (AoII) has recently been proposed as an alternative to existing information freshness metrics for real-time sampling and estimation problems involving information sources that are tracked by remote monitors. Different from existing metrics, AoII penalizes the incorrect information by increasing linearly with time as long as the source and the monitor are de-synchronized, and is reset when they are synchronized back. While AoII has generally been investigated for discrete time information sources, we develop a novel analytical model in this paper for push- and pull-based sampling and transmission of a continuous time Markov chain (CTMC) process. In the pull-based model, the sensor starts transmitting information on the observed CTMC only when a pull request from the monitor is received. On the other hand, in the push-based scenario, the sensor, being aware of the AoII process, samples and transmits when the AoII process exceeds a random threshold. The proposed analytical model for both scenarios is based on the construction of a discrete time MC (DTMC) making state transitions at the embedded epochs of synchronization points, using the theory of absorbing CTMCs, and in particular phase-type distributions. For a given sampling policy, analytical models to obtain the mean AoII and the average sampling rate are developed. Numerical results are presented to validate the analytical models as well as to provide insight on optimal sampling policies under sampling rate constraints.

I. INTRODUCTION

In real-time sampling and estimation problems involving information sources (or sensors) that are tracked by remote monitors, the performance of the estimator is indirectly quantified by a number of information freshness metrics, such as age of information (AoI) [1], [2] and metrics derived from it, e.g., age of incorrect information (AoII) [3], query-age of information [4], age penalty [5], etc. The most common feature of these metrics is that the penalty linearly increases with time while the information is stale at the monitor. In particular, AoII has been proposed in [3] to quantify information freshness by allowing AoII to increase only when there is a mismatch between the observed source and its estimator, and it can be reset back to zero even when this agreement is satisfied without having to take a new sample. It also differs from other metrics such as binary freshness [6] and MMSE (minimum mean square error) by further increasing the penalty when the estimation is kept incorrect for a long time.

This work is done when N. Akar is on sabbatical leave as a visiting professor at University of Maryland, MD, USA, which is supported in part by the Scientific and Technological Research Council of Türkiye (Tübitak) 2219-International Postdoctoral Research Fellowship Program.

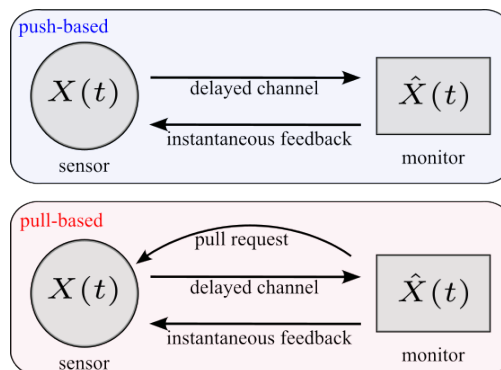


Fig. 1. Push-based and pull-based models for a sensor and a monitor.

AoII metric has generally been used in discrete time settings [3], [7]–[10], and a system model in continuous time is studied for the first time in this paper, to the best of our knowledge. Most practical systems are of continuous time nature which is the main reason for our choice of a continuous time setting. In [3], [7], [10], push-based transmission models are investigated, and an optimum threshold policy for a discrete time Markov chain (DTMC) source with homogeneous transition probabilities is proposed by using Markov decision processes (MDP). In [9], a pull-based transmission system for AoII is studied in which a monitor observes multiple homogeneous Markov sources with infinite state spaces. They provide an optimum threshold by using the Whittle index policy. Again, for the pull-based model, similar system models are investigated for various age-related metrics [5], [6], [11]. In [5] and [11], the sampling rates are allowed to change according to the latest estimation, similar to this paper. While binary freshness is used as the freshness metric in [11], an optimum policy to minimize age-penalty (defined as the time after the latest state transition) is derived in [5]. In [6], a remote monitor samples multiple CTMC sources under a sampling constraint in continuous time, and closed-form expressions for optimum sampling policies are derived for three freshness metrics which are fresh when equal, fresh when sampled, and fresh when close. Another recent work in a continuous time setting is [12], where the probability distribution of AoI, and the joint probability distribution of the observed process $X(t)$ and its estimation $\hat{X}(t)$, are derived.

In the pull-based model, the sensor starts transmitting information on the observed CTMC only when a pull request

from the monitor is received. In particular, the monitor is allowed to make pull requests according to a Poisson process with intensity dependent on the current estimation. In the push-based scenario, the sensor, being aware of the AoII process, samples and transmits when the AoII process exceeds a random threshold (in particular Erlang- k distributed), which depends on the current value of the estimator. Transmission times are random (exponentially distributed) in both models allowing us to study the impact of random network delays on the AoII process.

The proposed analytical models for both push- and pull-based scenarios are based on the construction of a DTMC making state transitions at the embedded epochs of synchronization points, using the theory of absorbing CTMCs, and in particular the *phase-type distribution* which stands for the distribution until absorption in a CTMC with one absorbing state [13]. The analytical models developed in this paper are based on the well-known properties of phase-type distributions which have also been used for information freshness and networked control problems in several existing works. In [14], the distribution of AoI and peak AoI metrics are derived by making use of phase-type distributions to deal with CTMCs with absorbing states. In another work [15], the phase-type distribution is used to find the expected time before a certain number of consecutive packet failures happens in a wireless closed-loop control system.

To summarize the contributions of this paper, i) we develop a novel and computationally efficient analytical model for obtaining the mean (or average) AoII in push- and pull-based sampling of CTMC processes, ii) the proposed analytical model relies on the well-known theory of absorbing CTMCs and can easily be extended to other settings including discrete time and other information freshness metrics, iii) it portrays a framework by which optimum parameters can be selected.

II. PRELIMINARIES

The following are preliminaries on absorbing Markov chains that are needed to follow the proposed analytical method. Consider a CTMC $Y(t), t \geq 0$, with K_1 transient and K_2 absorbing states, and \cdot . Let the generator of this CTMC be written as following,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (1)$$

where $\mathbf{A}_{K_1 \times K_1}$ and $\mathbf{B}_{K_1 \times K_2}$ correspond to the transition rates among the transient states, and from the transient states to the absorbing states, respectively. Upon merging all the absorbing states into one, the distribution of time until absorption, denoted by T , is known as the phase-type distribution [13] whose probability density function (PDF) is written in terms of the matrix exponential function of $t\mathbf{A}$ as,

$$f_T(t) = -\beta e^{t\mathbf{A}} \mathbf{A} \mathbf{1}, \quad (2)$$

from which the first and second moments can be obtained,

$$\mathbb{E}[T] = -\beta \mathbf{A}^{-1} \mathbf{1}, \quad \mathbb{E}[T^2] = 2\beta \mathbf{A}^{-2} \mathbf{1}, \quad (3)$$

where $\beta = \{\beta_i\}$ is a $1 \times K_1$ row vector with β_i standing for the initial probability in being transient state i , and $\mathbf{1}$ denotes a column vector of ones of appropriate size. Moreover, the probability of being absorbed in absorbing state j is,

$$p_j = -\beta \mathbf{A}^{-1} \mathbf{B} e^{(j)}, \quad (4)$$

where $e^{(j)}$ is a column vector of zeros of appropriate size except for its j th element which is one.

The absorbing CTMC whose generator is given in (1) can be converted to a DTMC with probability transition matrix as,

$$\begin{bmatrix} \mathbf{D} & \mathbf{E} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (5)$$

at the embedded epochs of state transitions. In particular, $\mathbf{D} = \{d_{ij}\}$ can be written as,

$$d_{ij} = \begin{cases} -\frac{a_{ij}}{a_{jj}}, & \text{if } i \neq j, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

A basic property about an absorbing Markov chain is that the expected number of visits to a transient state j starting from a transient state i is given by the (i, j) th entry of the so-called fundamental matrix $\mathbf{F} = (\mathbf{I} - \mathbf{D})^{-1}$ [16], where \mathbf{I} denotes the identity matrix of suitable size.

III. SYSTEM MODEL

Consider a sensor observing a finite-state irreducible CTMC process $X(t) \in \mathcal{N} = \{1, 2, \dots, N\}$ with generator matrix $\mathbf{Q} = \{q_{ij}\}$, where the state holding time at state i , denoted by H_i , is exponentially distributed with parameter $\sigma_i = -q_{ii}$, i.e., $H_i \sim \text{Exp}(\sigma_i)$ with $\mathbb{E}[H_i] = \frac{1}{\sigma_i}$, after which a state transition takes place to state j with transition probability $p_{ij} = -q_{ij}/q_{ii}$, for $i \neq j$. The monitor uses the last received update as its estimation, and the mismatch between the observed process and its estimation is measured with the AoII process, which is expressed as,

$$\text{AoII}(t) = t - \max(\tau | X(\tau) = \hat{X}(\tau), \tau \leq t). \quad (7)$$

We separately consider the push-based and pull-based transmission models which are illustrated in Fig. 1.

In the push-based model, the sensor decides when to start transmitting its observation to the monitor. We assume a channel with exponentially distributed service time with parameter μ to account for network delays. However, we assume instantaneous feedback from the monitor and therefore the sensor is always aware of the remote estimator and also of the AoII process maintained at the monitor. The sensor then waits for the AoII to exceed a random threshold Γ_i which depends on the current value i of the estimator. In this paper, we focus on Erlang- k distributed thresholds where an Erlang- k random variable $E \sim \text{Erlang-}k(\eta)$ with mean $\frac{1}{\eta}$ is the sum of k exponential random variables with the same mean $\frac{1}{k\eta}$, whose variance approaches zero as $k \rightarrow \infty$, i.e., deterministic distribution. Clearly, the exponential distribution is a sub-case of the Erlang- k distribution when $k = 1$.

In the pull-based model, sampling is initiated with a pull request from the monitor which sends these requests according to a Poisson process with intensity depending on the current value of the estimator. Pull requests from the monitor reach the sensor immediately, similar to instantaneous feedback. In both models, we assume that the sensor preempts its transmission if the observation changes before the transmission is over. After the preemption, while a new transmission is always initiated in the push-based model, a new pull request is required for initiating a transmission in the pull-based model. Additionally, a constraint on the sampling rate is enforced so that the sensor can not transmit the observations constantly. We leave two-way network delays, more advanced channel models (such as queued models), and deterministic thresholds for future work.

IV. ANALYTICAL METHOD

We first need the following definition: A time point t is a synchronization point (SP) at value S_i if $X(t)$ and $\hat{X}(t)$ are just synchronized at the value i , i.e., $X(t) = \hat{X}(t) = i$, $X(t^-) \neq \hat{X}(t^-)$. The interval between the SP t at value S_i and the next SP is called cycle- i . Subsequently, we consider a DTMC embedded at the epochs of synchronization points whose states are the synchronization values $S_i, i \in \mathcal{N}$. While the transition from S_i (at the current SP) to S_j (at the next SP) for $j \neq i$ is incurred with a timely reception of a sample, a self-transition from S_i to itself can also occur when $X(t)$ returns to its initial value i without a new sampling operation. Let p_{ij} and π_i denote the transition probability from synchronization value S_i to S_j , and the steady-state probability of being in state S_i at an SP. The steady-state vector $\pi = \{\pi_i\}$ satisfies

$$\pi = \pi P, \quad P\mathbf{1} = \mathbf{1}, \quad (8)$$

where $P = \{p_{ij}\}$. Then, we obtain π as,

$$\pi = \mathbf{1}^T (P + \mathbf{1}\mathbf{1}^T - I)^{-1}. \quad (9)$$

We note that the actual values of the transition probabilities will be dependent on the model (push- or pull-based) and the particular sampling policy of interest.

Let a_i , d_i , and r_i , denote the area under the AoII curve, duration, and average number of samples taken, respectively, of cycle- i . Let AoII and R denote the time-average of the AoII process and the average sampling rate of the overall system, respectively. Then,

$$\text{AoII} = \frac{\sum_{n=1}^N \pi_n a_n}{\sum_{n=1}^N \pi_n d_n}, \quad R = \frac{\sum_{n=1}^N \pi_n r_n}{\sum_{n=1}^N \pi_n d_n}. \quad (10)$$

In the next two subsections, for push- and pull-based sampling, respectively, and for a given sampling policy, we will first derive the transition matrix P of the embedded DTMC from which we will obtain π , and the quantities a_i , d_i , and r_i , $i \in \mathcal{N}$ which will allow us to obtain AoII and R via (10).

A. Push-Based Sampling

A sample path for the push-based sampling model is illustrated in Fig. 2 for $N = 3$ starting from S_1 . Let γ_1 be a realization of the random threshold Γ_1 . When AoII(t)

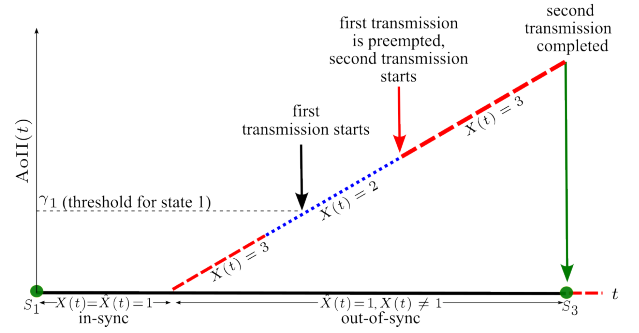


Fig. 2. A sample path for $X(t)$, $\hat{X}(t)$ and AoII(t) for push-based transmission for an example scenario. Green circles denote the synchronization points.

TABLE I
TRANSITION RATES FOR THE ABSORBING CTMC PROCESS $Y_i(t)$ FOR PUSH-BASED SAMPLING.

Transition Rates from (j, l) for $Y_i(t)$		
To	Condition	Value
$(j, l + 1)$	$l \neq k$	$k\theta_i$
(j', l)	$j' \neq j, \forall l$	$q_{jj'}$
S_i	$\forall l$	q_{ji}
S_j	$l = k$	μ

reaches γ_1 , the sensor samples $X(t)$ at state 2, and starts transmitting this observation which requires an exponentially distributed service time with parameter μ . However, before the transmission is complete, a state transition from state 2 to state 3 takes place. Therefore, the sensor preempts the previous transmission and re-samples $X(t)$ at state 3. Finally, the system reaches the next synchronization value S_3 at which $X(t)$ and $\hat{X}(t)$ are synchronized at 3, since this second transmission has completed without encountering a state change in the original process. It is also possible that we can reach the next synchronization value S_1 by a state change of the original process $X(t)$ from states 2 or 3 to the original state 1, but this situation is not illustrated in Fig. 2.

In the following, we focus on the particular synchronization value S_i , $i \in \mathcal{N}$ and describe the analytical method to find p_{ij} , $j \in \mathcal{N}$, a_i , d_i , and r_i when $\Gamma_i \sim \text{Erlang-}k(\theta_i)$. The procedure needs to be repeated $\forall S_j$, $j \neq i$. Let us revisit Fig. 2 and concentrate on the interval when AoII(t) > 0 until AoII(t) becomes zero. We will now show that this interval can be modeled as a phase-type distribution described in Section II with $Y_i(t)$ representing the corresponding absorbing CTMC process. For this purpose, we define a transient state as the pair (j, l) , where $j \neq i$ is the actual state of $X(t)$ and $l \in \{0, \dots, k - 1\}$ corresponds to the stage of the Erlang- k random variable and $l = k$ represents the situation that the threshold value is exceeded. In state (j, l) for $l < k$, the system is out-of-sync but the threshold is not reached yet, and the absorbing CTMC $Y_i(t)$ starts operation from state $(j, 0)$ with probability $\frac{q_{ij}}{\sigma_i}$. In state (j, k) , the system is still out-of-sync but the threshold is reached and therefore there is an ongoing transmission. Finally, we have the absorbing states S_j , $j \in \mathcal{N}$. According to the notation of Section II, we have

$K_1 = (k + 1)(N - 1)$ transient states and $K_2 = N - 1$ absorbing states. Then, the generator for $Y_i(t)$ is in the form (1) with the matrices \mathbf{A} , \mathbf{B} , and β of (1) replaced with $\mathbf{A}^{(i)}$, $\mathbf{B}^{(i)}$, and $\beta^{(i)}$, respectively, for which the characterizing matrix $\mathbf{A}^{(i)}$ can be written as,

$$\begin{bmatrix} \mathbf{Q}^{(-i)} - k\theta_i \mathbf{I} & k\theta_i \mathbf{I} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^{(-i)} - k\theta_i \mathbf{I} & k\theta_i \mathbf{I} & \dots & \mathbf{0} \\ \vdots & & & & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{Q}^{(-i)} - \mu \mathbf{I} & \end{bmatrix} \quad (11)$$

and the matrix $\mathbf{B}^{(i)}$ and the vector $\beta^{(i)}$ can be written as,

$$\mathbf{B} = \begin{bmatrix} \mathbf{q}^{(i)} & \mathbf{0} \\ \vdots & \vdots \\ \mathbf{q}^{(i)} & \mu \mathbf{I} \end{bmatrix}, \quad \beta^{(i)} = \begin{bmatrix} \frac{1}{\sigma_i} \mathbf{q}_r^{(i)} & \underbrace{0 \dots 0}_{k(N-1)} \end{bmatrix} \quad (12)$$

where $\mathbf{Q}^{(-i)}$ is \mathbf{Q} but its i th row and i th column removed, and $\mathbf{q}^{(i)}$ is the i th column of \mathbf{Q} excluding q_{ii} , and $\mathbf{q}_r^{(i)}$ is the i th row of \mathbf{Q} excluding q_{ii} .

We observe that the duration before absorption in this model, denoted by $T^{(i)}$, is equal to the duration of the out-of-sync interval in cycle- i . Since $T^{(i)}$ is phase-type distributed, the area under the AoII curve a_i and the duration d_i of cycle- i using (3) are,

$$a_i = \mathbb{E} \left[\frac{(T^{(i)})^2}{2} \right] = \beta^{(i)} (\mathbf{A}^{(i)})^{-2} \mathbf{1}, \quad (13)$$

$$d_i = \mathbb{E}[T^{(i)} + H_i] = -\beta^{(i)} (\mathbf{A}^{(i)})^{-1} \mathbf{1} + \frac{1}{\sigma_i}. \quad (14)$$

Finally, we can calculate from (4) the transition probability from the synchronization value S_i to S_j , p_{ij} , as

$$p_{ij} = -\beta^{(i)} (\mathbf{A}^{(i)})^{-1} \mathbf{B}^{(i)} \mathbf{e}^{(j)}. \quad (15)$$

In order to obtain r_i , we need to find the average number of times a transient state (j, k) is visited before absorption. For this purpose, we first obtain the absorbing MC in discrete time with the probability transition matrix in the form

$$\begin{bmatrix} \mathbf{D}^{(i)} & | & \mathbf{E}^{(i)} \\ \hline \mathbf{0} & | & \mathbf{0} \end{bmatrix} \quad (16)$$

with $\mathbf{D}^{(i)}$ obtained from $\mathbf{A}^{(i)}$ as in (6), resulting in the fundamental matrix $\mathbf{F}^{(i)}$,

$$\mathbf{F}^{(i)} = (\mathbf{I} - \mathbf{D}^{(i)})^{-1}. \quad (17)$$

Any time a state (j, k) is visited, a new sample is taken. Therefore, following Section II, we can write r_i as the average number of visits to any of the states (j, k) as,

$$r_i = \beta^{(i)} \mathbf{F}^{(i)} \mathbf{f}, \quad (18)$$

where \mathbf{f} is a column vector whose last $N - 1$ elements are one, and its remaining elements are zero. This concludes the calculations required for cycle- i . We repeat this procedure for all cycles to obtain the discrete time probability transition matrix \mathbf{P} using (15) and its steady-state vector π using (9),

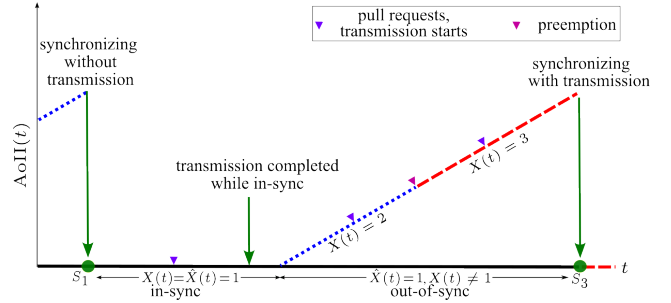


Fig. 3. A sample path for $X(t)$, $\hat{X}(t)$ and $\text{AoII}(t)$ for pull-based transmission for an example scenario.

TABLE II
TRANSITION RATES FOR THE ABSORBING CTMC PROCESS $Y_i(t)$ FOR PULL-BASED SAMPLING.

Transition Rates from (j, l) for $Y_i(t)$		
To	Condition	Value
$(j, 1)$	$l = 0$	λ_i
$(j', 0)$	$j' \neq j, l = 0, 1$	$q_{jj'}$
S_i	$l = 0, 1$	q_{ji}
S_j	$l = 1$	μ

which subsequently allow us to obtain AoII and the average sampling rate R through the expressions in (10), completing the analysis for $\Gamma_i \sim \text{Erlang-}k(\theta_i)$.

B. Pull-Based Sampling

In this scheme, the monitor sends pull requests according to a Poisson process with rate λ_i when the last estimation is $\hat{X}(t) = i$, and the sensor transmits the sample with an exponentially distributed service time with parameter μ . Different from the push-based model, if the observed process $X(t)$ changes before the transmission is completed and the sensor preempts the transmission, the sensor does not start another transmission until it receives another pull request. Further, in this model, a transmission may also happen while the source and the monitor are in-sync. However, this redundant sample does not affect the AoII and the estimation maintained at the monitor, but increases the average sampling rate.

A sample path of this transmission model is illustrated in Fig. 3 for an example scenario with $N = 3$, while pull-requests are illustrated with blue arrowheads, and preemption and completion times of transmissions are shown with purple and green arrowheads, respectively. At the beginning, the synchronization value S_1 is reached when $X(t)$ changes from state 2 to state 1, which matches with the estimation at this SP. While in-sync, a pull request is received upon which a transmission starts which is eventually completed. However, this particular transmission does not have an effect on the evolution of AoII. Synchronization is broken after $X(t)$ visits state 2 upon which two further pull requests are received. While the first transmission is preempted because of the state change, the transmission following the second pull request is completed at which the synchronization value S_3 is reached.

Similar to the push-based model, we introduce the absorbing CTMC process $Y_i(t)$ corresponding to the out-of-sync interval

in cycle- i . The CTMC $Y_i(t)$ has absorbing states $S_i, i \in \mathcal{N}$ and transient states as the pair (j, l) where $j \neq i$ corresponds to the state of $X(t)$, and l is a binary state $\{0, 1\}$. In state $(j, 0)$, the sensor waits for a pull request to start transmission and $Y_i(t)$ starts operation from this state with probability $\frac{q_{ij}}{\sigma_i}$, and state $(j, 1)$ indicates an ongoing transmission. All transition rates are summarized in Table II. By using them, we can express the characterizing matrices $\mathbf{A}^{(i)}$, $\mathbf{B}^{(i)}$ and $\beta^{(i)}$ of the absorbing CTMC $Y_i(t)$ as in Section II,

$$\mathbf{A}^{(i)} = \begin{bmatrix} \mathbf{Q}^{(-i)} - \lambda_i \mathbf{I} & \lambda_i \mathbf{I} \\ \mathbf{Q}^{(-i)} - \text{diag}(\mathbf{Q}^{(-i)}) & \text{diag}(\mathbf{Q}^{(-i)}) - \mu \mathbf{I} \end{bmatrix}, \quad (19)$$

$$\mathbf{B}^{(i)} = \begin{bmatrix} \mathbf{q}^{(i)} & \mathbf{0} \\ \mathbf{q}^{(i)} & \mu \mathbf{I} \end{bmatrix}, \quad \beta^{(i)} = \begin{bmatrix} \frac{1}{\sigma_i} \mathbf{q}_r^{(i)} & \underbrace{0 \ \cdots \ 0}_{N-1} \end{bmatrix}, \quad (20)$$

where $\text{diag}(\mathbf{M})$ is a diagonal matrix composed of the diagonal elements of \mathbf{M} . Again, a_i, d_i, p_{ij} are obtained using (13), (14) and (15), respectively. However, in order to calculate r_i , we also should consider the number of transmissions while in-sync, which is equal to the average number of pull requests before synchronization is broken. We can model that by considering an absorbing CTMC $Z_i(t)$ with two transient states and one absorbing state. In the first state (which we initially start the operation at), we wait for a pull request with rate λ_i whereas we complete transmission with rate μ in the second state. On the other hand, the absorbing state is reached upon a state change in the original process which occurs with rate σ_i . Therefore, the characterizing matrices $\hat{\mathbf{A}}^{(i)}$ and $\hat{\mathbf{D}}^{(i)}$ of the process $Z_i(t)$ given in Section II can be expressed as,

$$\hat{\mathbf{A}}^{(i)} = \begin{bmatrix} -\sigma_i - \lambda_i & \lambda_i \\ \mu & -\mu - \sigma_i \end{bmatrix}, \quad \hat{\mathbf{D}}^{(i)} = \begin{bmatrix} 0 & \frac{\lambda_i}{\lambda_i + \sigma_i} \\ \frac{\mu}{\mu + \sigma_i} & 0 \end{bmatrix}, \quad (21)$$

which results in the following expression for r_i ,

$$r_i = \beta^{(i)} (\mathbf{I} - \mathbf{D}^{(i)})^{-1} \mathbf{f} + [1 \ 0] (\mathbf{I} - \hat{\mathbf{D}}^{(i)})^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (22)$$

$$= \beta^{(i)} (\mathbf{I} - \mathbf{D}^{(i)})^{-1} \mathbf{f} + \frac{\lambda_i (\mu + \sigma_i)}{\lambda_i \sigma_i + \sigma_i \mu + \sigma_i^2}, \quad (23)$$

where \mathbf{f} is a column vector whose last $N - 1$ elements are one, and its remaining elements are zero. It is worth noting that the average sampling rate may be smaller than the overall pull request rate since the sensor ignores pull requests that arrive during an ongoing transmission.

V. NUMERICAL RESULTS

In the first numerical example, we verify the push-based analytical model with simulation results for three different scenarios. For all of these simulations, we select the channel transmission rate as $\mu = 1$ and $\mathbb{E}[\Gamma_i] = \theta^{-1}$. Three source processes are studied with the following generators,

$$\mathbf{Q}_1 = \{q_{1,ij}\}_{5 \times 5}, \quad q_{1,ij} = 0.25, \quad j \neq i, \quad q_{1,ii} = -1, \quad (24)$$

$$\mathbf{Q}_2 = \begin{bmatrix} -1 & 0.7 & 0.3 \\ 0.2 & -0.6 & 0.4 \\ 0.1 & 0.7 & -0.8 \end{bmatrix}, \quad \mathbf{Q}_3 = \begin{bmatrix} -0.5 & 0.5 \\ 0.75 & -0.75 \end{bmatrix}, \quad (25)$$

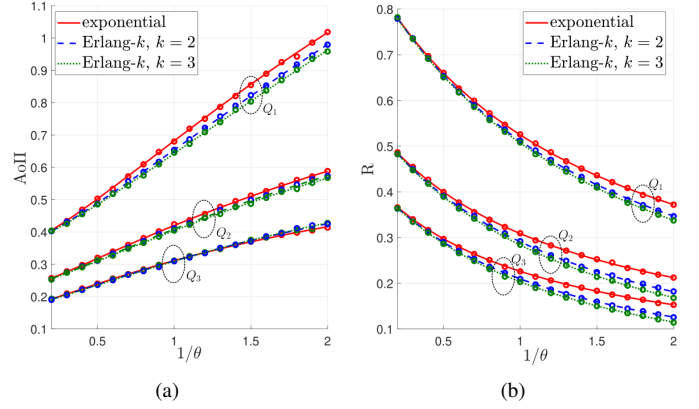


Fig. 4. a) AoII, b) average sampling rate R of the push-based transmission model as a function of θ^{-1} , for three CTMC processes and for three different values of the parameter k . While lines are for the analytical results, circles correspond to the simulation results.

where the first one has homogeneous states and the other two having heterogeneous states. In Fig. 4, AoII and average sampling rate R are depicted as a function of θ^{-1} for all the three information sources for Erlang- k distributed thresholds for $k = 1$ (i.e., exponentially distributed case), 2, 3 using both simulations and the analytical model. We observe an exact agreement between simulations and analytical results for all the cases. When θ^{-1} increases, we attain lower sampling rates at the expense of an increase in AoII. Additionally, we observe that when the order of the Erlang- k distribution increases, i.e., the variance of the threshold decreases, then a lower AoII is obtained at a lower average sampling rate. Studying Erlang- k distribution with larger k or the deterministic distribution, for improved AoII performance is left for future work.

Next, we focus on the information source with generator \mathbf{Q}_3 with $N = 2$, $\mathbb{E}[\Gamma_i] = \theta_i^{-1}, i = 1, 2$, and employ $k = 3$ while varying θ_1 and θ_2 . In Fig. 5, we study the performance of the sampling policies characterized with the pair $(\theta_1^{-1}, \theta_2^{-1})$, using a contour map. The red contour lines (resp. gray contour lines) correspond to the set of policies with the same average sampling rate (resp. same AoII). Additionally, for a given sampling rate budget, the policy that results in the minimum AoII is marked by a cross. We conclude that, if there is sufficient sampling rate budget, the best policy is to choose a small threshold for the faster state (state 2 with a lower state holding time), and use the rest of the budget for the slower state (state 1), and the best strategy converges to the $(0, 0)$ policy while the sampling rate budget increases.

In the second numerical example, similar simulations are repeated for the pull-based model with the same generators given in (24) and (25) and $\mu = 1$. In Fig. 6, AoII is depicted as a function of pull request rate $\lambda = \lambda_i, i \in \mathcal{N}$ both with simulations and the analytical model, and there is perfect agreement between the two curves. Moreover, we observe that AoII decreases and R increases monotonically when λ is increased. Additionally, while the rate of pull requests are lower than the transmission rate $\mu = 1$, almost all pull requests lead to a sampling operation resulting in a linear relation between them. As the rate of pull requests increases, some

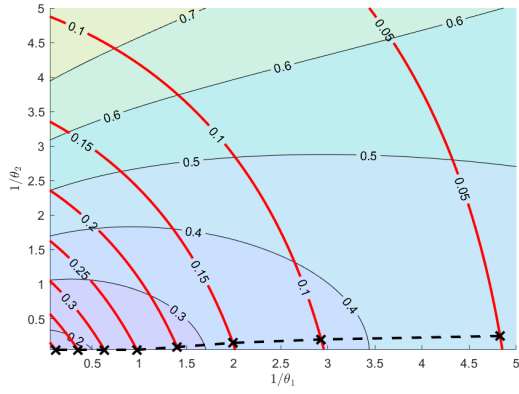


Fig. 5. Contour maps of AoII and R with respect to the sampling policy $(\theta_1^{-1}, \theta_2^{-1})$. The points which result in the minimum AoII values among the policies with the same average sampling rate pairs are marked with a cross.

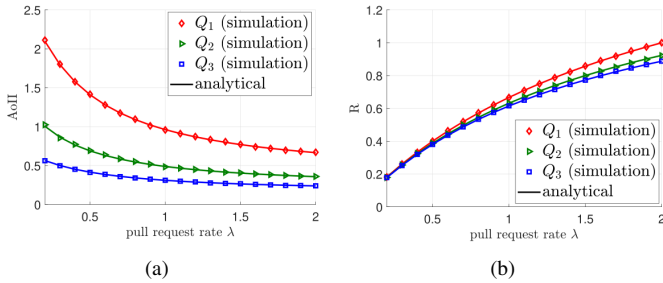


Fig. 6. a) AoII, b) average sampling rate R of the pull-based transmission model as a function of the common pull request rate λ .

of these requests arrive when a transmission is ongoing and they are therefore ignored. In Fig. 7, we study the performance of the sampling policies for the pull-based model characterized with the pair $(\lambda_1^{-1}, \lambda_2^{-1})$, using a contour map using the same convention as in Fig. 5. Unlike the push-based scenario, we observe that frequent sampling while the estimation is the slow state may not lead to the best policy under a given sampling rate budget. We observe that a more balanced sampling policy (note the positions of the cross markers in Fig. 7) is more well-suited for the pull-based model. The reason behind this is that while the push-based system only samples when the estimation is out-of-sync, the pull-based system may take samples even when in-sync, which increases the average sampling rate without impacting the AoII performance.

VI. CONCLUSIONS AND FUTURE WORK

We studied the average AoII and the average sampling rate of push- and pull-based sampling systems when the information source is a CTMC process. We developed novel and computationally efficient analytical models for both systems using a common framework based on the theory of absorbing CTMCs. For the push-based system and using the analytical models, we have observed that both AoII and average sampling rate performances are improved when the variance of the sampling threshold decreases. Moreover, the optimum sampling policy (leading to minimum AoII under a sampling budget) leads to lower sampling thresholds for slower states. We have also observed that this situation is very different in pull-

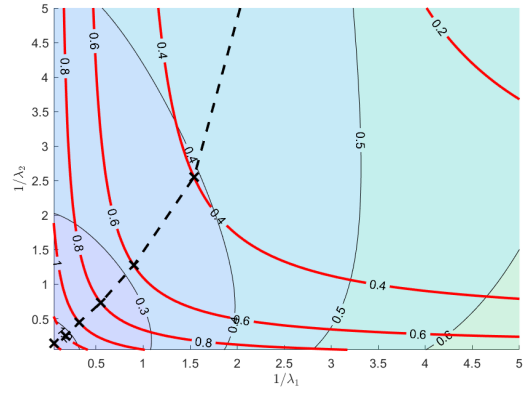


Fig. 7. Contour maps of AoII and R with respect to the sampling policy $(\lambda_1^{-1}, \lambda_2^{-1})$. The points which result in the minimum AoII values among the policies with the same average sampling rate pairs are marked with a cross.

based systems for which we have observed that more balanced pull requests lead to optimum sampling. Future work will consist of two-way delays, other sampling policies including deterministic thresholds in push-based systems, and the use of MDPs or reinforcement learning in deriving optimum policies.

REFERENCES

- [1] R. D. Yates and S. K. Kaul. The age of information: Real-time status updating by multiple sources. *IEEE Trans. Inf. Theory*, 65(3):1807–1827, March 2019.
- [2] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE J. Sel. Areas Commun.*, 39(5):1183–1210, May 2020.
- [3] A. Maatouk, S. Kriouile, M. Assaad, and A. Ephremides. The age of incorrect information: A new performance metric for status updates. *IEEE/ACM Trans. Netw.*, 28(5):2215–2228, October 2020.
- [4] F. Chiariotti, J. Holm, A. E. Kalør, B. Soret, S. K. Jensen, T. B. Pedersen, and P. Popovski. Query age of information: Freshness in pull-based communication. *IEEE Trans. Comm.*, 70(3):1606–1622, January 2022.
- [5] J. P. Champati, M. Skoglund, M. Jansson, and J. Gross. Detecting state transitions of a Markov source: Sampling frequency and age trade-off. *IEEE Trans. Comm.*, 70(5):3081–3095, March 2022.
- [6] N. Akar and S. Ulukus. Optimum monitoring of heterogeneous continuous time Markov chains. 2023. Available online at arXiv:2310.02223.
- [7] C. Kam, S. Kompella, and A. Ephremides. Age of incorrect information for remote estimation of a binary Markov source. In *IEEE Infocom*, July 2020.
- [8] S. Kriouile and M. Assaad. Minimizing the age of incorrect information for real-time tracking of Markov remote sources. In *IEEE ISIT*, July 2021.
- [9] S. Kriouile and M. Assaad. When to pull data from sensors for minimum distance-based age of incorrect information metric. February 2022.
- [10] Y. Chen and A. Ephremides. Minimizing age of incorrect information for unreliable channel with power constraint. In *IEEE Globecom*, December 2021.
- [11] M. Bastopcu and S. Ulukus. Timely tracking of infection status of individuals in a population. In *IEEE Infocom*, May 2021.
- [12] Y. Inoue and T. Takine. AoI perspective on the accuracy of monitoring systems for continuous-time Markovian sources. In *IEEE Infocom*, April 2019.
- [13] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. SIAM, 1999.
- [14] N. Akar and E. O. Gamgam. Distribution of age of information in status update systems with heterogeneous information sources: An absorbing Markov chain-based approach. *IEEE Commun. Lett.*, 27(8):2024–2028, May 2023.
- [15] L. Scheuven, T. Höbller, P. Schulz, N. Franchi, A. N. Barreto, and G. P. Fettweis. State-aware resource allocation for wireless closed-loop control systems. *IEEE Trans. Comm.*, 69(10):6604–6619, July 2021.
- [16] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer, 1960.