

# Detecting State Transitions of a Markov Source: Sampling Frequency and Age Trade-off

Jaya Prakash Champati, Mikael Skoglund, and James Gross

Information Science and Engineering, EECS, KTH Royal Institute of Technology, Stockholm, Sweden

E-mail: {jp, skoglund, jamesgr}@kth.se

**Abstract**—We consider a finite-state Discrete-Time Markov Chain (DTMC) source that can be sampled for detecting the events when the DTMC transits to a new state. Our goal is to study the trade-off between sampling frequency and staleness in detecting the events. We argue that, for the problem at hand, using Age of Information (AoI) for quantifying the staleness of a sample is conservative and therefore, introduce *age penalty* for this purpose. We study two optimization problems: minimizing average age penalty subject to an average sampling frequency constraint, and minimizing average sampling frequency subject to an average age penalty constraint; both are Constrained Markov Decision Problems. We solve them using a linear programming approach, and the resulting Markov policies are optimal among all causal policies. Our numerical results demonstrate that the computed Markov policies not only outperform optimal-periodic sampling policies, but also achieve sampling frequencies close to or lower than that of an optimal clairvoyant sampling policy by allowing a small age penalty.

## I. INTRODUCTION

Detecting the occurrence of an event when monitoring an information source or a process of interest is essential to applications from varied domains that include control and information systems. In a control system, for instance, a sensor samples a process for detecting events such as the state of the process exceeding a certain threshold. In an information system, a web crawling application is equipped with the task of downloading remote web pages to a local database (for page ranking/indexing etc.), and needs to detect the events when the remote web page gets updated.

In practice, it is impossible to know the exact time instant of occurrence of an event unless the source is sampled infinitely often or in every time slot possible. However, higher sampling frequency incurs costs to a system such as energy of a sensor used for sampling, or the bandwidth of the network used for transmitting the samples. On the other hand, using lower sampling frequency will result in staleness in detecting an event. Therefore, we are interested in the question: *given the source is sampled in a time slot  $n$ , how to choose the next sampling instant  $n + \tau$  such that the average sampling frequency and average staleness in the detection of the events are optimized?* In this work, we address this question for an information source modelled using a finite-state DTMC and the events of interest are transitions of the DTMC to new states. Even though this problem seems fundamental and arises in different applications, to the best of our knowledge, it has not been tackled in the literature; see Section V for related works.

The first step in studying the trade-off between sampling frequency and staleness is to choose an appropriate metric for quantifying the staleness of a sample. For this purpose, one may choose Age of Information (AoI), proposed in [1], that has emerged as a relevant performance metric for quantifying staleness of updates at a destination in a communication system. It is defined as the time elapsed since the generation of freshest update available at the destination. However, we argue that using AoI is conservative for the problem at hand and introduce a staleness metric *age penalty*, which is defined as the time elapsed since the first transition out of the most recently observed state. We then formulate two problems: minimizing average age penalty subject to an average sampling frequency constraint, and minimizing average sampling frequency subject to an average age penalty constraint. Both the problems are Constrained Markov Decision Problems (CMDPs) and we use Linear Programming (LP) approach to solve for policies that are optimal among all causal policies. In our numerical analysis, using a two-state Markov chain, we demonstrate the superior performance of the computed policies to that of the optimal-periodic sampling policies. We also present a comparison of the sampling frequency achieved by the optimal policy with that of the sampling frequency of an optimal clairvoyant (non-causal) sampling policy.

The rest of the paper is organized as follows. In Section II, we present the system model and the formulations of the two optimization problems. The CMDP solution for both the problems is described in Section III. We present numerical analysis for two-state Markov chain in Section IV and related work in Section V. Finally, we conclude in Section VI.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

### A. Markov Source

We consider an information source/process that is modelled by an  $N$ -state DTMC  $\{X_n, n \geq 0\}$  where  $N < \infty$ . We assume that the DTMC is ergodic, i.e., irreducible and aperiodic. Let  $S = \{1, 2, \dots, N\}$  denote the set of states. We use  $p_{ij}$ , for all  $i, j \in S$ , to denote the one-step probabilities, and the  $n$ -step transition probabilities are denoted by

$$p_{ij}^{(n)} = \mathbb{P}(X_n = i | X_0 = j), \forall i, j \in S.$$

Given the one-step probabilities, the  $n$ -step transition probabilities can be computed using matrix multiplication on the one-step transition probability matrix [2]. Let  $\xi_j$  denote the stationary probability of finding the DTMC in state  $j$ .

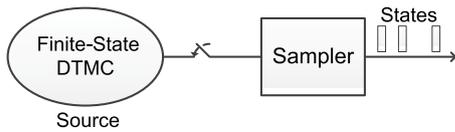


Fig. 1: Sampling an information source/process modelled using a DTMC. Each sample reveals the state of the DTMC.

A time slot in the system represents one unit of time of the DTMC and the state transitions occur at the start of a time slot. The state of the DTMC can only be observed by sampling the source; see Figure 1. Let  $T_0 = 0, T_1, T_2, \dots$  denote the time instants of transitions of the DTMC to new states. We are interested in detecting these transitions at the earliest time possible. Our motivation for studying this problem arises due to its relevance to applications from different domains.

- In a control system, the source is a process of interest, and a state transition represents an event where the process exceeds a certain threshold.
- In a web crawling application [3], the source is a remote web page and the state transitions models the updating events of the website.

Clearly, sampling the source at the start of every slot allows us to detect each and every transition of the DTMC. Instead, our aim here is to use lower sampling frequency. This translates to energy savings for a sensor and/or bandwidth savings for transmitting lower number samples to a controller/monitor. In the case of web crawling application, this translates to lower frequency of downloads of the remote web page. However, using lower sampling frequency will result in staleness in detecting a transition and may also miss several transitions. We are thus interested in studying the trade-off between sampling frequency and staleness. Next, we define sampling policies and present our definition of age penalty for quantifying the staleness.

### B. Sampling Policy and Age Penalty

Assume that  $X_0$  is given. A sampling policy  $\pi$  specifies the set of sampling instants  $\{G_k, k \geq 1\}$ , where  $G_k$  is the sampling instant of the  $k$ th sample. Define  $\tau_k = G_k - G_{k-1}$  for all  $k \geq 1$ , and  $G_0 = 0$ , then the policy  $\pi$  can be equivalently specified by  $\{\tau_k, k \geq 1\}$ . We assume that  $\tau \in Q = \{1, 2, \dots, M\}$ , where  $M < \infty$  is the maximum inter-sampling time allowed in the system. Let  $\Pi$  denote the set of all causal policies, where a causal policy considers the current and all past observed states and past actions for choosing the current action. In the sequel, we study the following policies.

- 1) *Markov policies*: A Markov policy maps each state to an action with a fixed probability. To be precise, let  $j$  be the observed state in the  $k$ th decision epoch, then under a Markov policy  $\tau_k$  is assigned a value  $\tau \in A$  according to a fixed probability distribution  $\mathbb{P}^\pi(\tau_k = \tau | j)$ . Let  $\Pi^{\text{MR}}$  denote the set of Markov policies.

- 2) *Periodic sampling policies*: Under these policies, samples are taken at fixed time intervals  $\tau$ . With a slight abuse in notation we use  $\pi(\tau)$  to denote such a policy. Note that periodic sampling policies are a subclass of Markov policies.
- 3) *Optimal clairvoyant sampling policy*: Under this policy, the next transition to a new state is assumed to be known a priori, and thus the source is sampled exactly at the instants when transitions between states occur. Let  $\pi^\dagger = \{G_k^\dagger, k \geq 1\}$  denote this policy and  $\nu^\dagger$  denote its average sampling frequency. Note that  $\pi^\dagger$  is a non-causal policy and we study it for theoretical benchmarking.

As stated before, sampling the source at the start of every slot allows us to identify each and every transition of the DTMC to a new state and thus staleness of each sample is zero. However, quantifying the staleness of a sample in general is not entirely obvious. This is because, when the sampler samples the source it may find that the DTMC is in the same state or a different state from the previous sample, and even in the former case multiple transitions might have occurred. One may consider AoI, denoted by  $\Delta(t)$ , at the sampler as the staleness metric. It increases linearly between two sampling instants and resets to zero at the sampling instants. However, using this statelessness metric is conservative in this context. To illustrate this, in Figure 2 we plot the sample-path of a 3-state DTMC and the resulting AoI. Note that in the duration between the instants  $G_1$  and  $G_2$ , DTMC stays in state 2 for 3 time-slots after it was observed by the sampler at  $G_1$ . Ideally, this should not be accounted for the staleness of the sample at  $G_2$ , but AoI adds a linear penalty for this duration.

Using the above insight, we quantify the statelessness of a sample  $k$  by introducing *age penalty*  $A_k$ , which is defined as the time elapsed since the first transition out of the state in the  $k-1$  sample<sup>1</sup>. Under policy  $\pi$  the age penalty for the  $k$ th sample is given by

$$A_k(\pi) = G_k - \min_n \{T_n : T_n \geq G_{k-1}\}.$$

This entity is illustrated and contrasted with the AoI in Figure 2. Under a policy  $\pi$ , the average age penalty  $\mathbb{E}[A(\pi)]$  is given by,

$$\mathbb{E}[A(\pi)] = \limsup_{K \rightarrow \infty} \frac{\mathbb{E}[\sum_{k=1}^K A_k(\pi)]}{K},$$

and the average sampling-interval is given by  $\limsup_{K \rightarrow \infty} \frac{\mathbb{E}[\sum_{k=1}^K \tau_k]}{K}$ , where the expectation is taken with respect to the probability distribution induced by  $\pi$  on the sequence of observed states and actions.

### C. Optimization problems $\mathcal{P}_1$ and $\mathcal{P}_2$

We are interested in the following problems. For a given upper bound  $\nu \in (0, 1]$  on the average sampling frequency,

<sup>1</sup>One may also consider including the number of missed transitions in the age penalty and with some effort solve the problem using the same approach in this paper.

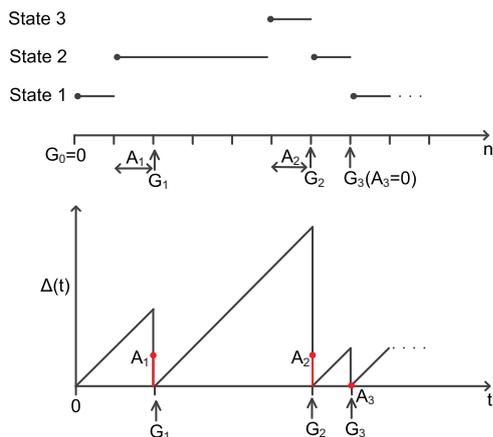


Fig. 2: A sample path of a 3-state Markov chain. AoI and age penalties are depicted for first three sampling instants of a policy with  $G_1 = 2, G_2 = 6$ , and  $G_3 = 7$ .

in problem  $\mathcal{P}_1$  we aim to minimize the average age penalty which is stated below.

$$\begin{aligned} & \underset{\pi \in \Pi}{\text{minimize}} && \mathbb{E}[A(\pi)] \\ & \text{s.t.} && \limsup_{K \rightarrow \infty} \frac{\mathbb{E}[\sum_{k=1}^K \tau_k]}{K} \geq \frac{1}{\nu}. \end{aligned} \quad (1)$$

For a given upper bound  $d \geq 0$  on the average age penalty, in problem  $\mathcal{P}_2$  we aim to maximize the average sampling-interval which is state below.

$$\begin{aligned} & \underset{\pi \in \Pi}{\text{maximize}} && \limsup_{K \rightarrow \infty} \frac{\mathbb{E}[\sum_{k=1}^K \tau_k]}{K} \\ & \text{s.t.} && \mathbb{E}[A(\pi)] \leq d. \end{aligned} \quad (2)$$

Let  $\pi_1^*$  and  $\pi_2^*$  denote optimal policies for  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , respectively.

**Remark:** For  $\mathcal{P}_1$ , an optimal-periodic sampling policy chooses  $\tau = \lceil 1/\nu \rceil$ . For  $\mathcal{P}_2$  an optimal-periodic sampling policy chooses  $\tau = d + 1$ .

Finally, we define  $\tau^\dagger = \lceil 1/\nu^\dagger \rceil$ .

### III. LINEAR PROGRAMMING SOLUTION APPROACH

Both  $\mathcal{P}_1$  and  $\mathcal{P}_2$  are Constrained Markov Decision Problems (CMDP). A CMDP with finite state and action sets has an optimal policy in the set of Markov policies [4], and it can be efficiently solved using the Linear Programming (LP) approach presented in [5]. Therefore, in the following we only consider the set of Markov policies. Under Markov policies the induced stochastic process  $\{X_{G_k}, k \geq 1\}$ , i.e., the sequence of observed states, is also a DTMC; in the sequel we refer to it as *induced DTMC*.

#### A. Elements of the CMDP

The decision epochs in  $\mathcal{P}$  are indexed by  $k$ .

- **State space:**  $S = \{1, 2, \dots, N\}$ .
- **Action space:** At decision epoch  $k$ , the next inter-sampling time  $\tau_{k+1}$  is chosen from the set  $Q = \{1, 2, \dots, M\}$ .

- **Transition probabilities:** The next state  $i \in S$  of the induced DTMC only depends on the current observed/sampled state  $j$  and the sampling interval  $\tau$ . To be precise, let  $i$  be the state observed in decision epoch  $k$ , i.e., in time slot  $G_k$ , then the transition probability of the induced DTMC to state  $j$  for any sampling interval  $\tau$  is given by

$$\begin{aligned} q_{j\tau i} &= \mathbb{P}(X_{G_k+\tau} = i | X_{G_k} = j) \\ &= \mathbb{P}(X_\tau = i | X_0 = j) \\ &= p_{ij}^{(\tau)}, \forall i, j \in S \text{ and } \tau \in Q. \end{aligned}$$

Further, given  $\pi \in \Pi^{\text{MR}}$ , the stationary probabilities  $\mathbb{P}^\pi(X_{G_k} = j)$  for the induced DTMC can be computed from the following transition probabilities.

$$\begin{aligned} \mathbb{P}(X_{G_{k+1}} = i | X_{G_k} = j) &= \mathbb{E}[q_{j\tau i}] \\ &= \sum_{\tau=1}^M q_{j\tau i} \mathbb{P}^\pi(\tau | j), \forall i, j \in S. \end{aligned} \quad (3)$$

- **Costs:** In decision epoch  $k$ , if the state is  $j$ , then choosing a sampling interval  $\tau \in A$  results in a cost contributing to the average age-penalty which is given by

$$c_{j\tau} = \sum_{n=1}^{\tau-1} (\tau - n)(1 - p_{jj})p_{jj}^{n-1},$$

and the cost contributing to the average sampling-interval is given by  $\tau$ . Note that  $c_{j\tau}$  is the expected number of slots the DTMC has spent after moving out of state  $j$  in the sampling interval  $\tau$ . It is easy to see that

$$\mathbb{E}[A_{k+1} | X_{G_k} = j, \tau_k = \tau] = c_{j\tau}, \forall k \geq 1 \quad (4)$$

#### B. LP formulations for $\mathcal{P}_1$ and $\mathcal{P}_2$

We define  $z_{j\tau}^\pi = \lim_{k \rightarrow \infty} \mathbb{P}(X_{G_k} = j, \tau_k = \tau)$ , the steady-state probability of observing the state-action pair  $(j, \tau)$  under a policy  $\pi \in \Pi^{\text{MR}}$ . Then, from (4), we obtain

$$\begin{aligned} \mathbb{E}[\Delta(\pi)] &= \sum_{j=1}^N \sum_{\tau=1}^M c_{j\tau} z_{j\tau}^\pi \\ \limsup_{K \rightarrow \infty} \frac{\mathbb{E}[\sum_{k=1}^K \tau_k]}{K} &= \sum_{j=1}^N \sum_{\tau=1}^M \tau z_{j\tau}^\pi. \end{aligned}$$

In the LP formulations for  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , we solve for  $z_{j\tau}^\pi$  with the following constraints,

$$\sum_{j=1}^N \sum_{\tau=1}^M z_{j\tau}^\pi = 1, \quad (5)$$

$$\sum_{\tau=1}^M z_{i\tau}^\pi = \sum_{j=1}^N \sum_{\tau=1}^M q_{j\tau i} z_{j\tau}^\pi, \quad i \in S, \quad (6)$$

$$z_{j\tau}^\pi \geq 0, \quad j \in S \text{ and } \tau \in A. \quad (7)$$

The constraint (6) is a consequence of the equilibrium equations for the induced DTMC in the steady-state. In the following, we present an equivalent LP formulation for  $\mathcal{P}_1$ ,

$$\begin{aligned} & \underset{\{z_{j\tau}^*\}}{\text{minimize}} && \sum_{j=1}^N \sum_{\tau=1}^M c_{j\tau} z_{j\tau}^* \\ & \text{s.t.} && \sum_{j=1}^N \sum_{\tau=1}^M \tau z_{j\tau}^* \geq \frac{1}{\nu}, \\ & && (5), (6), (7). \end{aligned} \quad (8)$$

Let  $\{z_{j\tau}^*\}$  denote the optimal solution for (8), then the stationary probabilities under  $\pi_1^*$  are computed as follows. For  $\tau \in A$ ,

$$\mathbb{P}^{\pi_1^*}(\tau|j) = \frac{z_{j\tau}^*}{\sum_{i=1}^N z_{i\tau}^*}, \quad j \in S.$$

Similarly, an equivalent LP can be formulated for  $\mathcal{P}_2$  and  $\pi_2^*$  can be obtained.

### C. Computing $\nu^\dagger$

Note that, in  $\mathcal{P}_1$  the value of  $\nu$  in the constraint can be chosen in the interval  $(0, 1]$ . We are particularly interested in setting  $\nu = \nu^\dagger$ , because this will give us the minimum achievable average age-penalty for the same sampling frequency achieved by the optimal clairvoyant sampling policy  $\pi^\dagger$ . We note that  $\nu^\dagger$  can be obtained by subtracting the percentage of the total frequency of transitions in the DTMC contributed due to self transitions, i.e., transitions from a state to itself, from the total frequency of transitions in the DTMC. Since a transition occurs in every time slot, total frequency of transitions in the DTMC is 1. The percentage of the total frequency of transitions in the DTMC contributed due to self transitions is given by  $\sum_{j=1}^N \xi_j p_{jj}$ . The following proposition follows directly from the above analysis.

**Proposition 1.** *Under the optimal clairvoyant sampling policy  $\pi^\dagger$ , the average sampling frequency  $\nu^\dagger$  is given by*

$$\nu^\dagger = 1 - \sum_{j=1}^N \xi_j p_{jj}.$$

For a two-state Markov chain, the steady-state probabilities are given by

$$\xi_1 = \frac{p_{21}}{p_{12} + p_{21}}, \quad \text{and} \quad \xi_2 = \frac{p_{12}}{p_{12} + p_{21}},$$

and

$$\nu^\dagger = \xi_1 p_{12} + \xi_2 p_{21} = \frac{2p_{12}p_{21}}{p_{12} + p_{21}}.$$

In Figure 3, we plot  $\nu^\dagger$  by varying  $p_{21}$  for different values of  $p_{12}$ .

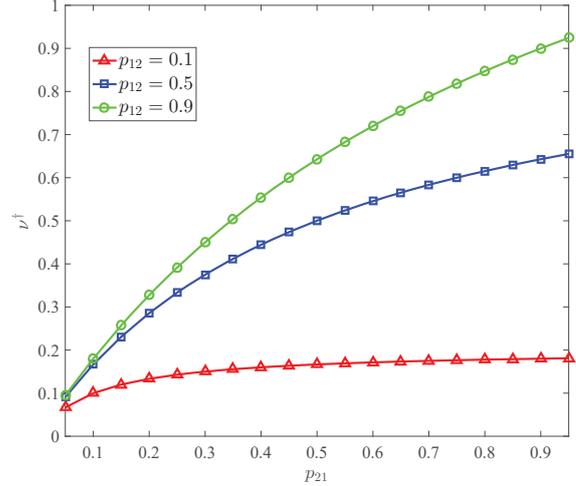


Fig. 3: Sampling frequency under  $\pi^\dagger$  for a two-state Markov chain.

## IV. NUMERICAL RESULTS: TWO-STATE MARKOV CHAIN

In this section we present numerical analysis for a two-state DTMC. Even though this is the simplest case, it can potentially be used in modelling sources where the set of states can be divided into two sets, for example “disturbance” vs “no disturbance”, and the events of interest are transitions between these sets. We have implemented the LPs using *linprog* in MATLAB. In the following, we first present two numerical examples to examine the structure of the optimal Markov policies for  $\mathcal{P}_1$  and  $\mathcal{P}_2$ . We then present sampling frequency and age penalty trade-off and a performance comparison between optimal and optimal-periodic sampling policies.

*Example 1:* In this example, we solve  $\mathcal{P}_1$  for transition probabilities  $p_{12} = 0.1$  and  $p_{21} = 0.6$ , and the constraint on the expected sampling interval is equal to  $1/\nu^\dagger = 5.83$ . The computation of the optimal policy  $\pi_1^*$  results in the following stationary probabilities,

$$\begin{aligned} \mathbb{P}^{\pi_1^*}(\tau = 6|j = 1) &= 0.465 \quad \text{and} \quad \mathbb{P}^{\pi_1^*}(\tau = 7|j = 1) = 0.535, \\ \mathbb{P}^{\pi_1^*}(\tau = 2|j = 2) &= 1. \end{aligned}$$

Note that the transition probability out of state 2 is higher and thus the policy sets  $\tau = 2$  whenever the observed state is 2. The minimum expected age penalty is computed to be 1.416. An optimal periodic sampling policy chooses  $\tau = \tau^\dagger = \lceil 1/\nu^\dagger \rceil = 6$ .

*Example 2:* In this example, we solve  $\mathcal{P}_2$  when  $p_{12} = 0.9$ ,  $p_{21} = 0.9$ , and the expected age penalty is upper bounded by  $d = 1$ . The computation of the optimal policy  $\pi_2^*$  results in the following stationary probabilities.

$$\mathbb{P}^{\pi_2^*}(\tau = 2|j) = 0.899 \quad \text{and} \quad \mathbb{P}^{\pi_2^*}(\tau = 3|j) = 0.101 \quad \text{for } j = 1, 2.$$

The minimum expected sampling frequency is computed to be 0.476. The optimal periodic sampling policy chooses  $\tau = d + 1 = 2$ , and hence the sampling frequency under this policy is 0.5.

### A. Performance Comparison

In Figure 4, we compare the average age penalties achieved by optimal-periodic sampler and the optimal policy  $\pi_1^*$  obtained by solving  $\mathcal{P}_1$  under the constraint  $\nu = \nu^\dagger$ . Recall that for this case, the optimal-periodic sampler sets the sampling interval equal to  $\tau^\dagger = \lceil 1/\nu^\dagger \rceil$ . From the figure, we observe that for lower transition probabilities between the states, i.e., lower  $p_{12}$  and  $p_{21}$  values, periodic sampler achieves age penalties only slightly higher than that of the optimal policy, because in this case the optimal policy is also choosing sampling intervals close to that of the periodic sampler. The gap between them, however, increases significantly for higher transition probabilities. The zigzag pattern of the periodic sampler can be attributed to the ceil function used in computing the sampling interval.

In Figures 5, and 6 we compare average sampling frequencies achieved by the optimal-periodic sampler and the optimal policy  $\pi_2^*$  by solving  $\mathcal{P}_2$ . From Figure 5, we observe the trade-off between achievable sampling frequencies and age penalties. As expected, for age penalty constraint of one time slot, i.e.  $d = 1$ , the achievable sampling frequency is lower than 0.5 for both policies. However,  $\pi_2^*$  results in much lower sampling frequencies for lower transition probabilities. In Figure 6, we set  $d = 1$  and thus the optimal-periodic sampler samples every 2 time slots with sampling frequency 0.5. On the other hand,  $\pi_2^*$  provides much lower sampling frequencies when either of the transition probabilities are small.

Finally, in Figure 7, we present the ratio between the expected sampling frequency achieved by  $\pi_2^*$  and  $\nu^\dagger$ , under average age penalty constraint  $d = 1$ . We note that under  $\pi^\dagger$  the age penalty is always zero. This cannot be achieved by any causal policy with a sampling frequency strictly less than one. Nonetheless, an interesting observation from the figure is that by allowing a small age penalty  $d = 1$ , the optimal policy  $\pi_2^*$  can achieve lower sampling frequency than  $\nu^\dagger$  when transition probabilities are higher, say  $p_{12} = 0.9$  and  $p_{21} = 0.9$ . For lower transition probabilities  $p_{12} = 0.9$  and  $p_{21} = 0.9$ , the ratio is always greater than 1, i.e., optimal policy  $\pi_2^*$  couldn't achieve the sampling frequency  $\nu^\dagger$  and may require more relaxation in the age penalty constraint. In conclusion, for lower transition probabilities, i.e., if the events become rare, the optimal policy performs worse with respect optimal clairvoyant sampling policy.

## V. RELATED WORKS

In this section we present the most related works<sup>2</sup> to the problem we studied.

In the AoI literature, the works [6]–[8] considered remote estimation of the states of a Markov Chain. However, all these works propose and use a different metric from ours for measuring the staleness of the updates and/or study specific

<sup>2</sup>Since the problem naturally arises and is inherent to several applications from varied domains, the related works present here are only representative.

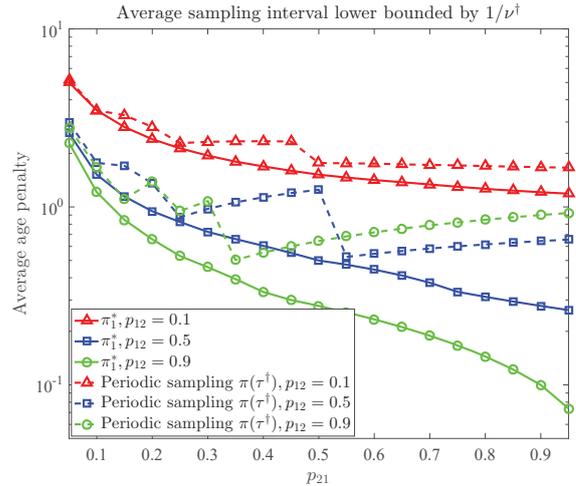


Fig. 4: Expected age penalties achieved by  $\pi_1^*$  and the optimal-periodic sampler for different  $p_{12}$  and  $p_{21}$  values.

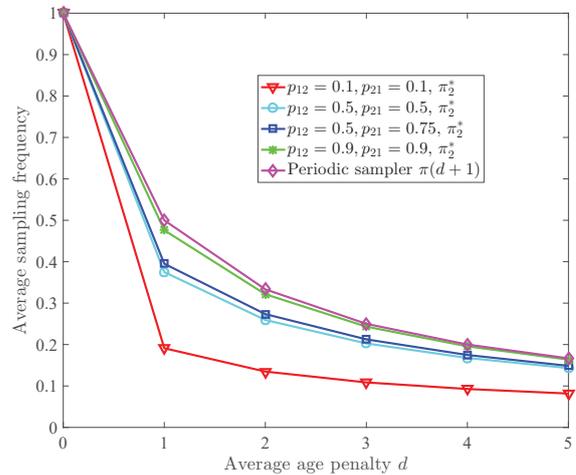


Fig. 5: Sampling frequency vs age penalty trade-off: expected sampling frequencies achieved by  $\pi_2^*$  and the optimal-periodic sampler by varying age penalty constraint value.

class of sampling policies. The problem of when to sample next has been studied for many years in control theory, see for example [9]–[12]. In [9] ([10]), the authors considered the off-line (on-line) problem of choosing the time instants to sample sensor measurements to minimize a Linear Quadratic Gaussian (LQG) cost in a Linear Time Invariant (LTI) system. In [11], the authors considered minimizing squared error distortion for state estimation of a Markov source under a constraint on maximum number of transmitted samples. We note however that, in this work the sensor is assumed to samples the process continuously but only transmits certain samples based on some criterion (event-triggering). In [12], the authors studied the design of sampling intervals such that the stability of a non-linear stochastic dynamical system is ensured.

Perhaps the most relevant application to the problem we have studied is the web crawling application [3], [13]. The

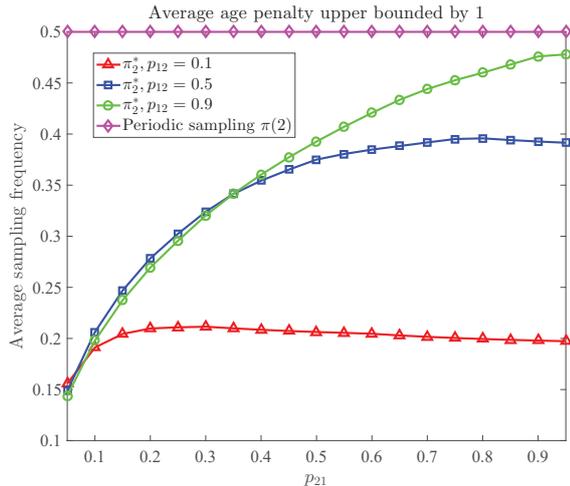


Fig. 6: Expected sampling frequencies achieved by  $\pi_2^*$  and the optimal-periodic sampler for different  $p_{12}$  and  $p_{21}$  values, and  $d = 1$ .

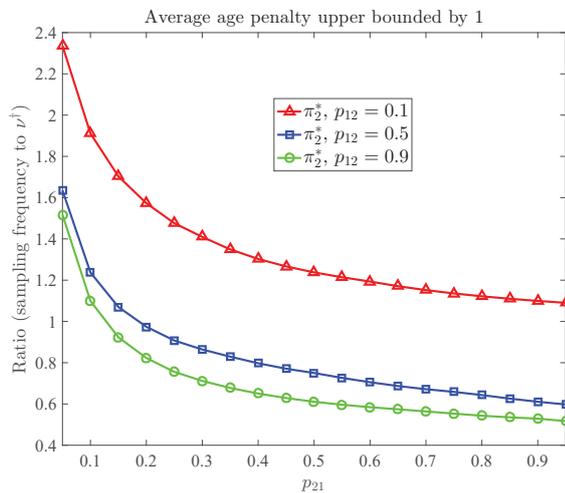


Fig. 7: Ratio between the expected sampling frequency achieved by  $\pi_2^*$  and  $\nu^\dagger$ .

authors in [13] have solved a static optimization problem for computing optimal fixed intervals between downloads for different web pages. To the best of our knowledge, dynamic policies that use the state of the system have not been studied in this line of work; see [3] for a survey. In contrast to above works, we considered the set of all causal sampling policies and studied the fundamental trade-off between sampling frequency and age penalty for detecting state transitions in a finite-state DTMC.

## VI. CONCLUSION

We have studied the trade-off between sampling frequency and staleness for detecting transitions of a DTMC to new states. The staleness of the  $k$ th sample is quantified using age penalty, which is defined as the time elapsed since the first transition out of the state in the  $k - 1$  sample. The formulated

problems  $\mathcal{P}_1$  and  $\mathcal{P}_2$  are CMDPs and were solved by deriving equivalent LPs. We have provided a closed-form expression for  $\nu^\dagger$ , the sampling frequency under the optimal clairvoyant sampling policy. Even though our problem setting looked simple, the numerical examples revealed that the optimal policies have randomized Markov policy structure, i.e., simple deterministic optimal policies may not exist for this problem. Apart from the superior performance of the computed optimal policy over optimal-periodic sampling policy, we found that by allowing a small age penalty the optimal policy achieves sampling frequency lower than  $\nu^\dagger$  in some cases.

We leave comprehensive simulation results considering  $N > 2$  for future work. We would like to explore different age penalties and study the trade-off when there are multiple sources. Finally, we are interested in studying the problem for different models for the information source.

## REFERENCES

- [1] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Proc. IEEE SECON*, 2011.
- [2] J. R. Norris, *Markov Chains*. Cambridge University Press, 1997.
- [3] C. Olston and M. Najork, "Web crawling," *Found. Trends Inf. Retr.*, vol. 4, no. 3, p. 175246, Mar. 2010.
- [4] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall, 1999.
- [5] A. S. Manne, "Linear programming and sequential decisions," *Management Science*, vol. 6, no. 3, pp. 259–267, 1960.
- [6] C. Kam, S. Kompella, G. D. Nguyen, J. E. Wieselthier, and A. Ephremides, "Towards an effective age of information: Remote estimation of a markov source," in *IEEE INFOCOM WKSHPs*, April 2018, pp. 367–372.
- [7] Y. Sun and B. Cyr, "Information aging through queues: A mutual information perspective," *CoRR*, vol. abs/1806.06243, 2018.
- [8] Y. Inoue and T. Takine, "Aoi perspective on the accuracy of monitoring systems for continuous-time markovian sources," in *IEEE INFOCOM WKSHPs*, April 2019, pp. 183–188.
- [9] H. Kushner, "On the optimum timing of observations for linear control systems with unknown initial state," *IEEE Transactions on Automatic Control*, vol. 9, no. 2, pp. 144–150, April 1964.
- [10] E. Skafidas and A. Nerode, "Optimal measurement scheduling in linear quadratic gaussian control problems," in *Proceedings of the 1998 IEEE International Conference on Control Applications (Cat. No. 98CH36104)*, vol. 2, Sep. 1998, pp. 1225–1229 vol.2.
- [11] M. Rabi, G. V. Moustakides, and J. S. Baras, "Adaptive sampling for linear state estimation," *SIAM Journal on Control and Optimization*, vol. 50, no. 2, pp. 672–702, 2012.
- [12] R. P. Anderson, D. Milutinovi, and D. V. Dimarogonas, "Self-triggered sampling for second-moment stability of state-feedback controlled sde systems," *Automatica*, vol. 54, pp. 8 – 15, 2015.
- [13] J. Cho and H. Garcia-Molina, "Effective page refresh policies for web crawlers," *ACM Trans. Database Syst.*, vol. 28, no. 4, p. 390426, Dec. 2003. [Online]. Available: <https://doi.org/10.1145/958942.958945>