2021 19th International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt)

Minimization of Age of Incorrect Estimates of Autoregressive Markov Processes

Bhavya Joshi*, Rajshekhar V Bhat†, B. N. Bharath‡ and Rahul Vaze§
*†‡Indian Institute of Technology Dharwad, India
§ Tata Institute of Fundamental Research, Mumbai, India
Email: *170020010@iitdh.ac.in, †rajshekhar.bhat@iitdh.ac.in, ‡bharathbn@iitdh.ac.in, §rahul.vaze@gmail.com

***Abstract*—We consider a source that sends freshness-sensitive status updates about an auto-regressive Markov process to a monitor, in a time-slotted system. The source samples the random process at the start of every slot and decides whether to transmit the sample or not. The transmission of a sample incurs a fixed cost. When a monitor receives a sample, its information about the source is perfect. However, when no samples are received, it estimates the realization of the process based on previously received samples. We adopt a metric referred to as the age of incorrect estimates (AoIE), defined as the product of an estimation error, E, and, v, the time elapsed since the latest time at which the monitor had a *sufficiently* correct estimate. We formulate an optimization problem to decide when a source must transmit a packet for minimizing the long-term average expected weighted sum of the AoIE and the transmission cost. We cast this problem as a Markov decision process and prove that the optimal policy is a *threshold*-type policy, in which, for a fixed v, there exists a threshold on E beyond which it is optimal to transmit, and vice versa. Using numerical simulations, we illustrate this threshold structure of the optimal policy. We also consider a simple periodic policy in which the information packets are transmitted periodically, after every fixed number of slots, irrespective of the realizations of E and v, and numerically show that its performance is significantly worse than that of the optimal threshold-type policy.**

I. Introduction

In many emerging Internet of Things (IoT) applications, sensing and delivery of information in a timely manner is extremely important [1]–[4]. A widely used metric for characterizing the timeliness of information is the age of information (AoI), defined as the *time* elapsed since the generation (at the source) of the last successful update packet delivered (to the monitor); a lower age implies a fresher information [5], [6].

A major shortcoming of this metric is that it does not account for the information content of the packets. Due to this, the instantaneous AoI linearly increases in the interval between two consecutive packet receptions, even if the monitor has a very accurate knowledge of the information content at the source, for instance, by estimation. Intuitively, (i) if the monitor has a *good* estimate of the realization of the random process, its information may not be considered *aged*, even if the monitor has not received an update packet recently, and (ii) even if a packet is received recently, if the estimate of the information content is inaccurate, it is sensible to consider it as an *aged* packet.

Based on this intuition, in order to account for (i) and (ii), a new metric, referred to as the age of incorrect information (AoII), has been proposed in [1]. Specifically, the instantaneous AoII at time t is defined as, $\Delta_{\text{AoII}}(t) \triangleq g(t) \times h(X(t), \hat{X}(t))$, where $g(t)$ is an age penalty (having units of time) for having incorrect information at the monitor, and $\hat{X}(t)$ is the estimate of $X(t)$ and $h(X(t), \hat{X}(t))$ is a quantification of its incorrectness. Several other modifications to the original AoI metric, such as age of synchronization [7] and age of changed information [8], have also been suggested as relevant metrics for characterizing the staleness of information in various scenarios.

Recently, there has also been an increased interest in characterizing staleness of remote estimates of various classes of information sources. The paper [9] ([10]) considered remote estimation of a Wiener process (an Ornstein-Uhlenbeck process, the continuous analogue of the first-order auto-regressive process) and showed that if the sampler does not have any knowledge of the signal being sampled, the optimal sampling strategy that minimizes a linear (non-linear) function of the original AoI metric is also optimal in minimizing the mean-squared estimation error. Among the works which consider the AoII metric, [1] and [2] consider information sources to be finite-state discrete-time Markov chains.

In this work, we consider a source that samples a discrete-time autoregressive Markov process in every slot. The source then decides whether to transmit the sample or not, to the monitor over a noiseless channel. In case no packet is delivered to the monitor, the monitor estimates the realization of the random process given the past received samples. At time t, let $E(t)$ be the squared estimation error at the monitor and $v_\epsilon(t)$ be the time elapsed since $E(t)$ was less than or equal to ϵ. Then, we define the instantaneous age of incorrect estimates (AoIE) at time t to be $\Delta_\epsilon(t) \triangleq v_\epsilon(t)E(t)$. Note that this is an instance of a more general AoII metric (proposed in [1]) mentioned above, with $g(t) = v_\epsilon(t)$ and $h(t) = E(t)$.[1] This instance of the AoII metric, i.e., the AoIE metric, penalises information

[1]Although this instance of the AoII metric is mentioned in [1], it is not analyzed in [1].

ISBN 978-3-903176-37-9 © 2021 IFIP

staleness scaled by a factor that accounts for estimation error. We consider the problem of minimizing the sum of long-term average AoIE and transmission cost and cast it as a Markov decision process (MDP).

The current work is related to [1]–[3] in the sense that all the works account for the information content in the AoI metric and adopt the MDP framework for obtaining their results. Specifically, current work and [1]–[3] adopt some form of the general AoII metric proposed in [1] and show that the optimal actions (the decision on whether to transmit an information packet or not) exhibit threshold structures in state variables of the MDPs considered. However, our work is significantly different from [1]–[3] due to the differences in the model adopted for the information process, the estimation strategy and the definition of correctness of an estimate, as described below.

- The current work considers a first order auto-regressive process with a noise having zero mean and finite variance with unimodal symmetric distribution function, as in [11], which forms a countably infinite-state Markov chain. However, [1]–[3] consider finite-state Markov chain sources, where the probabilities of remaining in the current state and transitioning to different states do not depend on the current state. Due to this, the information process considered by the current paper is more general.
- For quantification and analysis of the difference in the information content of the process, $X(t)$ and its estimate $\hat{X}(t)$, we adopt a squared error metric, $E(t) = (X(t) - \hat{X}(t))^2$, along with a least squares estimation strategy. Whereas, [1], [2] assume $E(t) = 0$ if $X(t) = \hat{X}(t)$ and $E(t) = 1$ otherwise, and [3] considers $E(t) = |X(t) - \hat{X}(t)|$ as the error metric. In [1]–[3], the latest successful packet delivered is considered to be the estimate of the process until receiving the next packet. Under a special case when the parameter of the autoregressive process is unity, our least squares estimation strategy specializes to that adopted in [1]–[3].
- Moreover, we deem an estimate to be correct when $E(t)$ is sufficiently small, which is controlled using a parameter, $\epsilon > 0$. However, [1]–[3] deem an estimation error to be correct, only when $E(t) = 0$.

Due to the above differences, obtaining the results in the paper is challenging and do not follow from [1]–[3] even when the communication channel is noiseless and reliable. Hence, unlike in [1]–[3] where the channel is assumed to be noiseless but unreliable, we consider that the channel is completely reliable.

The main contribution of the paper is: *we prove that the optimal policy for minimizing the sum of long-term average AoIE and transmission cost has a threshold structure in E (the estimation error), for a fixed v_ϵ (the time elapsed since the estimation error, E was less than ϵ) and vice versa*. The result is obtained by showing that the state-action value function is submodular, which is the challenging part of the work. The result has two important implications: First, the knowledge of the threshold structure of the optimal policy can be exploited to reduce computational complexity of the value iteration algorithm that we use for obtaining the optimal policy and computation of the thresholds. Second, real-time implementation of the transmitter controller is simple as we only need to store the thresholds and compare the instantaneous states with them for executing the optimal policy.

II. System Model

We consider a source sending updates to a monitor in the form of information packets. The system is time-slotted with slot index denoted by $t \in \mathbb{N}$. Let the discrete-time random process being observed at the source be denoted by $X(t)$ for $t \in \{1, 2, \dots\}$.[2] At time instant t, the source can either transmit the information sample $X(t)$[3] to the monitor, incurring a unit cost, or stay idle. The corresponding received signal at the monitor is given by $Y(t) = X(t)I(t)$, where $I(t)$ is defined as

$$I(t) = \begin{cases} 1 & \text{if the source sends an update in slot } t, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

In other words, for the sake of simplicity, the channel between the monitor and the source is assumed to be noiseless. In case the source does not send an update, the monitor estimates the current sample based on all the previously received samples as follows:

$$\hat{X}(t) = \begin{cases} X_t & \text{if } I(t) = 1, \\ f\left(X_{\{k<t:I(k)=1\}}\right) & \text{otherwise,} \end{cases} \quad (2)$$

where $X_{\{k<t:I(k)=1\}}$ is the set of all samples delivered before the time instant t, and f is the estimation strategy used. A detailed description of the estimation strategy is provided in Sec. II-B. We define the instantaneous estimation error at the monitor as

$$E(t) = (X(t) - \hat{X}(t))^2. \quad (3)$$

Whenever $E(t) \leq \epsilon$, we say that the monitor has the *correct estimate* of the data at the source at time t. Consequently, the instantaneous age of incorrect estimates (AoIE) is defined as

$$\Delta_\epsilon(t) = v_\epsilon(t)E(t), \quad (4)$$

where $v_\epsilon(t)$ is the time elapsed since the latest time the monitor had the correct estimate. Note that $v_\epsilon(t) = t - \mu_\epsilon(t)$, where $\mu_\epsilon(t)$ is the latest time at which the monitor had a correct estimate of the packet, i.e.,

$$\mu_\epsilon(t) = \max\{\tau \leq t : E(\tau) \leq \epsilon\}. \quad (5)$$

In the next section, we present the problem formulation along with the model for the evolution of $X(t)$.

[2]The model for $X(t)$ is described in Sec. II-B.

[3]In a more general case, the source may transmit some function of all the previously observed samples. However, for the source we consider, transmitting the current sample is sufficient.

ISBN 978-3-903176-37-9 © 2021 IFIP

A. Problem Formulation

Our goal in this work is to obtain a stationary deterministic[4] policy π that minimizes the long-term average expected weighted sum of AoIE and the cost of transmission, for a given estimator, f. Here, the policy, $\pi = \{I(t)\}_{t=0,1,\ldots}$ provides a rule to decide whether the sensor should send information to the monitor or not in a slot. Mathematically, we are interested in solving the following stochastic optimization problem:

$$J(s) = \min_{\pi} \lim_{T\to\infty} \frac{1}{T}\left(\sum_{t=1}^{T} \mathbb{E}_{\pi}\left[\Delta_{\epsilon}(t) + \lambda I(t)\right]\right), \tag{6}$$

where $\lambda > 0$ is a fixed real number, that trades-off the relative importance of the AoIE and transmission cost, and $\mathbb{E}_{\pi}$ denotes expectation with respect to the randomness in $X(t)$ and possible randomness in the policy π.

Remark 1. *At any slot t, the source can compute the instantaneous estimation error. This is because it samples[5] $X(t)$ at the start of slot t, and it can also estimate $\hat{X}(t)$ based on the previous samples it has transmitted, using the estimator f. Therefore, based on $X(t)$ and $\hat{X}(t)$, the source can compute $\mu_{\epsilon}(t)$, the latest time instant at which the monitor had the correct estimate, using which it can obtain $\Delta_{\epsilon}(t)$. Hence, the source can solve the optimization problem in* (6)*. Since the monitor does not know the true realization of $X(t)$ when $I(t) = 0$, it cannot compute $\Delta_{\epsilon}(t)$.*

In order to solve (6), for concreteness, we consider an autoregressive information evolution model along with a least squares estimation strategy, which we describe below.

B. Source Model and Estimation Strategy

We consider a simple auto-regressive process model for $X(t)$, which evolves as follows:

$$X(t) = aX(t-1) + W_t, \; t = 0, 1, \ldots, \tag{7}$$

where $a \in \mathbb{R}$ is any fixed and known constant, and $X(-1) := 0$. For simplicity of presentation, we assume that $W_t \in \mathcal{W}$, $t \in \mathbb{N}$ are independent and identically distributed (i.i.d.) *discrete* random variables, where W_t satisfies the following properties:

i. The distribution function, $p_{W_t}(\cdot)$ is symmetric, i.e., $p_{W_t}(-x) = p_{W_t}(x)$ for all $x \in \mathcal{W}$.
ii. W_t is unimodal, which satisfies

$$\mathbb{P}\{x - \epsilon \le W_t \le x + \epsilon\} \le \mathbb{P}\{y - \epsilon \le W_t \le y + \epsilon\}, \tag{8}$$

for any $x > y > 0$ and $\epsilon > 0$.

Many practical noise distributions, including the widely adopted Gaussian noise, admit the above properties. It is also important to note that all our results can be easily extended to the case when W_t, $t \in \mathbb{N}$ are i.i.d. continuous random variables with the above properties.

Next, we describe the estimation strategy adopted in this work. At time t, suppose that the latest sample received by the monitor is X_{t-L}, received at time slot $t-L$, for some positive integer L. Using the fact that $X(t)$ evolves as (7), we can write the following:

$$X(t) = a^L X(t-L) + \sum_{i=0}^{L} a^i W_{t-i}. \tag{9}$$

Based on this, at time t, the least squares estimate of $X(t)$ at the monitor given the sample $X(t-L)$, is given by

$$\hat{X}(t) = a^L X(t-L).$$

In summary, under the considered auto-regressive source model and the least squares[6] estimation strategy, (2) can be specialized to

$$\hat{X}(t) := \begin{cases} X(t) & \text{if } I(t) = 1, \\ a^L X(t-L) & \text{otherwise}, \end{cases} \tag{10}$$

where $t - L$ is the time instant of the latest successful packet delivery. It is important to note the distinction between L and v_{ϵ}; L is the time elapsed since the latest successful packet delivery, whereas v_{ϵ} is the time elapsed since the latest time the monitor had a "correct" estimate.

In the following section, considering the above model and the estimation procedure, we prove that a *threshold-type* policy is optimal for (6).

III. Markov Decision Process (MDP)-Based Solution

In order to prove the main result of this paper, in the following subsection, we show that the problem in (6) is an instance of an MDP.

A. Formulation

Towards casting (6) as an MDP, in the following, we identify the state space, action space and the associated cost.

1) State Space: At time t, the state of the system is $s_t := (E(t), v_{\epsilon}(t))$, where $v_{\epsilon}(t) = t - \mu_{\epsilon}(t)$. Since we have considered that the noise process, W_t, is discrete, $E(t)$ takes values from a countable subset of non-negative real

[4] In a later section of the paper, we cast the problem being formulated as an infinite horizon discounted cost MDP with countably infinite states and binary actions. Using standard techniques, the MDP can be shown to have a stationary deterministic policy as the optimal policy, as in [12, Lemma 6], [13, Theorem 6.2.7]. However, for brevity and ease of presentation, we directly consider the class of stationary randomized policies.

[5] We assume that sampling $X(t)$ costs zero units of energy, as in [11].

[6] One could investigate other estimation strategies at the expense of notational complexity.

ISBN 978-3-903176-37-9 © 2021 IFIP

values, $\mathbb{R}^+$ and $v_\epsilon(t) \in \mathbb{Z}^+$. Hence, the state space of the system is a countable (infinite) subset of $\mathbb{R}^+ \times \mathbb{Z}^+$.

2) Action Space: The action space is denoted by $\mathcal{A} := \{0,1\}$, and the corresponding decision variable at time slot t is $I(t)$.
3) Instantaneous Cost: The instantaneous cost (possibly random due to the randomness in the process and the transmission decisions) in slot t is $C_{\epsilon,t} := \Delta_\epsilon(t) + \lambda I(t)$.

Given the above setup, it is clear that the state transition probabilities and cost values depend only on the current state $S_t = s$, and not on the entire history of states and actions taken. In principle, the state transition probabilities can be computed using (3), (4), (5) and (7). Therefore, the problem in (6) can be written as an MDP. In what follows, instead of considering (6) as an average cost problem, we will consider minimizing the infinite horizon discounted cost MDP with a discount factor of $0 < \gamma < 1$. This is because, it can be shown that the discounted cost MDP converges to the average cost MDP, and it is easier to derive the structure of the optimal policy to (6) using the discounted cost MDP [12].

Given an initial state $s = (E, v)$, the optimal total expected γ-discounted cost corresponding to (6) is given by

$$J_\gamma(s) = \min_\pi \lim_{T\to\infty} \left(\sum_{t=1}^{T} \gamma^t \mathbb{E}_\pi \left[\Delta_\epsilon(t) + \lambda I(t) | \, s \right] \right). \quad (11)$$

We can now write down the following discounted cost optimality equation:

$$J_\gamma(s) = \min_{I\in\{0,1\}} \left(\mathbb{E}[\Delta_\epsilon + \lambda I | s, I] + \gamma \mathbb{E}[J_\gamma(s')|s, I] \right), \quad (12)$$

and define the following value iteration:

$$J_{\gamma,t+1}(s) = \min_{I\in\{0,1\}} \left(\mathbb{E}[\Delta_\epsilon + \lambda I | s, I] + \gamma \mathbb{E}[J_{\gamma,t}(s')|s, I] \right). \quad (13)$$

It can be shown that $J_{\gamma,t+1}(s) \to J_\gamma(s)$ as $t \to \infty$ for every s and γ [12]. Moreover, the discounted expected cost in (6), $J_\gamma(s)$ converges to the average expected cost, $J(s)$, as $\gamma \to 1$. Concretely, $\lim_{\gamma\to1}(1-\gamma)J_\gamma(s) \to J(s)$ for any state s.

This result can be proved by verifying a set of technical conditions along the lines in [14]. Hence, we consider the discounted expected cost criterion in the rest of the work. We also drop the subscript ϵ from Δ_ϵ and v_ϵ for notational convenience.

B. Threshold Structure of the Optimal Policy

In this subsection, we show that the optimal policy that solves (6) has a threshold structure, by considering the discounted cost criteria in (11). The value function, $J_\gamma(E, v)$ for $s = (E, v)$ in (12) can be expressed as

$$J_\gamma(E, v) = \min_I Q_\gamma(E, v, I), \quad (14)$$

where the state-action value function (Q-function), $Q_\gamma(E, v, I)$, is defined as

$$Q_\gamma(E, v, I) := \\ \mathbb{E}\left[\Delta' + \lambda I | E, v, I\right] + \gamma \mathbb{E}_{E',v'}\left[J_\gamma(E', v')|E, v, I\right]. \quad (15)$$

In the above, the parameter $0 < \gamma < 1$ is the discount factor and $I \in \{0,1\}$. The optimal policy for solving (12) in state (E, v) is given by

$$I^*(E, v) = \operatorname*{argmin}_{I\in\{0,1\}} Q_\gamma(E, v, I). \quad (16)$$

In the following, for a fixed v, we prove that the optimal $I^*(E, v)$ solving (12) has a threshold structure in E, where $I^*(E, v) = 1$ beyond a certain threshold on E. For proving this, it is sufficient to prove that $I^*(E+e, v) \geq I^*(E, v)$ where $e \in \mathbb{R}^+$, i.e., $I^*(\cdot)$ is monotonic. This follows from the submodularity[7] of $Q_\gamma(E, v, I)$ in I and E [15], which we prove in the following. Subsequently, a similar threshold structure of $I^*(E, v)$ on v for a fixed E is also proven.

1) Threshold Structure of I in E for a Fixed v: The following result proves the submodularity of the Q-function in E and I, and the threshold structure of the optimal policy, I, in E, for a fixed value of v.

Theorem 1. *For a fixed v, the Q-function, $Q_\gamma(E, v, I)$ is submodular, i.e.,*

$$Q_\gamma(E+e, v, 1) - Q_\gamma(E+e, v, 0) \leq \\ Q_\gamma(E, v, 1) - Q_\gamma(E, v, 0), \quad (17)$$

and I has a threshold structure in E for a fixed v.

Proof. We begin by proving that if no transmission occurs, the expected instantaneous cost incurred starting at a state with higher estimation error is not less than that when started at a state with lower estimation error. This is shown in the following lemma, which is proved in Appendix A:

Lemma 2. *For any v, E and $e > 0$, $\mathbb{E}_{\Delta'}\left[\Delta'|E+e, v, 0\right] \geq \mathbb{E}_{\Delta''}\left[\Delta''|E, v, 0\right]$.*

Similarly, we have the following lemma, which is proved in Appendix B, which says that the expected infinite horizon cost starting at a state with a higher estimation error is not less than that when we start at a state with a lower estimation error, when no transmission occurs.

Lemma 3. *For a fixed v, the following inequality on the average value function holds $\mathbb{E}\left[J_\gamma(E', v')|E+e, v, 0\right] \geq \mathbb{E}\left[J_\gamma(E'', v'')|E, v, 0\right]$.*

[7] Suppose a function $g(\cdot,\cdot)$ is such that $g(x, u+1) - g(x, u)$ is decreasing in x, $g(\cdot,\cdot)$ is a submodular function [15].

ISBN 978-3-903176-37-9 © 2021 IFIP

Using Lemma 2 and Lemma 3, we prove Theorem 1 as follows. Consider

$$\begin{aligned}
&Q_\gamma(E+e,v,1) - Q_\gamma(E+e,v,0) \\
&= \big(0+\lambda+\gamma J_\gamma(0)\big) - \big(\mathbb{E}_{\Delta'}\left[\Delta'+0|E+e,v,0\right] \\
&\qquad\qquad +\gamma\mathbb{E}\left[J_\gamma(E',v')|E+e,v,0\right]\big) \\
&\overset{(a)}{\leq} \big(0+\lambda+\gamma J_\gamma(0)\big) - \big(\mathbb{E}_{\Delta''}\left[\Delta''+0|E,v,0\right] \\
&\qquad\qquad +\gamma\mathbb{E}\left[J_\gamma(E'',v'')|E,v,0\right]\big) \\
&= Q_\gamma(E,v,1) - Q_\gamma(E,v,0),
\end{aligned} \tag{18}$$

where (a) follows from Lemma 2 and Lemma 3. Hence, (17) holds, and from Theorem 3.1.1 in [15], it follows that I is monotonic in E. Hence, I has a threshold structure in E for a fixed v. ☐

Similar to Theorem 1, we next show that I has a threshold structure in v for a fixed E.

2) Threshold Structure of I in v for a Fixed E: We consider E to be fixed. We have the following result:

Theorem 4. *The following holds:*

$$\begin{aligned}
&Q_\gamma(E,v+1,1) - Q_\gamma(E,v+1,0) \\
&\leq Q_\gamma(E,v,1) - Q_\gamma(E,v,0).
\end{aligned} \tag{19}$$

and I has a threshold structure in v for a fixed E.

Proof. For proving the theorem, we need the following results:

Lemma 5. $\mathbb{E}_{\Delta'}\left[\Delta'+0|E,v+1,0\right] = \mathbb{E}_{\Delta''}\left[\Delta''+0|E,v,0\right]$.

Lemma 6. *For a fixed E, the value function* $\mathbb{E}\left[J_\gamma(E',v')|E,v+1,0\right] \geq \mathbb{E}\left[J_\gamma(E'',v'')|E,v,0\right]$.

Lemma 5 and Lemma 6, respectively, show that if no transmission occurs, the expected instantaneous and infinite horizon costs incurred starting at a state with a higher v are not less than those when started at a state with a lower v. Lemma 5 is proved in Appendix C and Lemma 6 can be proved along the lines in that of Lemma 3.

Now consider

$$\begin{aligned}
&Q_\gamma(E,v+1,1) - Q_\gamma(E,v+1,0) \\
&= \big(0+\lambda+\gamma J_\gamma(0)\big) - \\
&\quad \big(\mathbb{E}_{\Delta'}\left[\Delta'+0|E,v+1,0\right] + \gamma\mathbb{E}\left[J_\gamma(E',v')|E,v+1,0\right]\big) \\
&\overset{(a)}{\leq} \big(0+\lambda+\gamma J_\gamma(0)\big) - \\
&\quad \big(\mathbb{E}_{\Delta''}\left[\Delta''+0|E,v,0\right] + \gamma\mathbb{E}\left[J_\gamma(E'',v'')|E,v,0\right]\big) \\
&= Q_\gamma(E,v,1) - Q_\gamma(E,v,0),
\end{aligned} \tag{20}$$

where (a) follows from Lemma 5 and Lemma 6. Based on the above discussion, as in the proof of Theorem 1, we conclude that I has a threshold structure in v for a fixed E. ☐

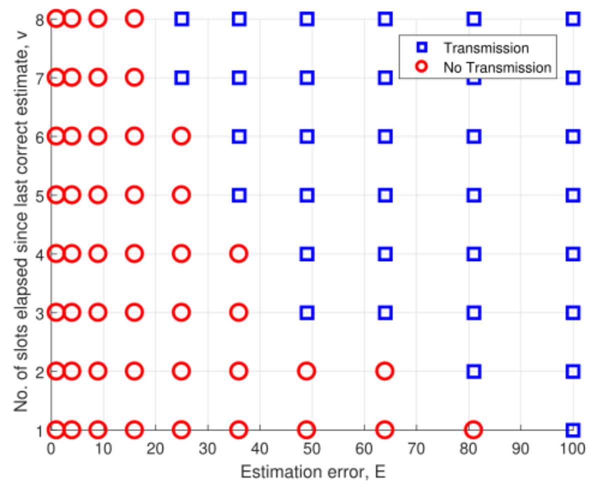

Fig. 1: Optimal actions for different states, (E, v), where E is the squared estimation error and v is the time elapsed since the last correct estimate. We observe that the optimal actions have a threshold structure in the states. That is, for a fixed E, there exists a threshold on v beyond which it is optimal to transmit and vice versa. We let $a = 1$, $\gamma = 0.9999$, $\lambda = 100$ and $\epsilon = 0$. W is uniformly distributed over $\{-1, 1\}$. In the figure, the markers are shown only for realizable E and v values, which are unevenly spaced for the chosen parameters.

Based on Theorem 1 and Theorem 4, we conclude that for a fixed v, there exists a threshold, E_v^{th}, on E, which depends on v, beyond which it is optimal to transmit. Similarly, for a fixed E, there exists a threshold, v_E^{th}, on v, which depends on E, beyond which it is optimal to transmit. Then, the optimal policy, which we refer to as the *threshold policy*, is the following: *For any realization of E and v in any slot, if $E \geq E_v^{\text{th}}$ or $v \geq v_E^{\text{th}}$, it is optimal to sample the information source and transmit.* The result has the following important implications:

- First, the knowledge of threshold structure can be exploited to reduce computational complexity in obtaining the optimal policy and computation of the thresholds, E_v^{th} and v_E^{th}, as elaborated below. We adopt the value iteration presented in (13) for obtaining the optimal policy and computation of the thresholds. In (13), for each iteration, t, one needs to compute the realization of the objective function for each state $s = (E, v)$ and then find the optimal I that minimizes the objective function for the state. The knowledge of the threshold structure of the optimal policy can be exploited to reduce the number of computations in the value iteration algorithm. Concretely, for some state $s = (E, v)$, let $I = 1$ be optimal. Then, due to the threshold structure, one can say that $I = 1$ is optimal for any state (E, v') and (E', v), where $v' > v$ and $E' > E$ without computing the objective functions for these states and optimizing them. When the value iteration in (13) converges (i.e., as $t \to \infty$), the optimal thresholds can be obtained from the optimal actions for each state. Concretely, the optimal threshold on E for a fixed v, $E_v^{\text{th}} = \min\{E : I^*(E, v) = 1\}$, where $I^*(E, v)$ is the optimal action in the state (E, v) presented in

ISBN 978-3-903176-37-9 © 2021 IFIP

(16). Similarly, the optimal threshold on v for a fixed E, $v_E^{\text{th}} = \min\{v : I^*(E, v) = 1\}$.

- Second, due to the threshold structure of the optimal solution, real-time implementation of the transmitter controller is simple as we only need to store the thresholds and compare the instantaneous states with them for executing the optimal policy. In absence of any knowledge of such structural properties, one would need to store the optimal actions for all the states, leading to higher storage capacity requirements.

IV. Simulation Results

In this section, we provide simulation results. The parameters adopted for obtaining the figures are mentioned in respective captions.

Recall from Section III-B that the optimal policy that solves (6) has the threshold structure in E for a given v and vice versa, where E is the estimation error and v is the time elapsed since the latest time the monitor had a correct estimate. We illustrate this threshold structure of the optimal policy in Fig. 1. From the figure, note that when $E = 49$, it is optimal to transmit whenever $v \geq 3$. Similarly, when $v = 2$, it is optimal to transmit for any $E \geq 81$. This indicates that $v_{49}^{\text{th}} = 3$ and $E_2^{\text{th}} = 81$. The thresholds v_E^{th} and E_v^{th} for other values of E and v, respectively, can be similarly computed.

In Fig. 2, we plot the average expected AoIE, Δ, against the average expected transmission cost (the first and second terms in (6), respectively). We benchmark the optimal threshold policy, which we obtain from value iteration in (13), with a naive periodic policy in which transmissions occur after every K slots, where K is chosen such that the threshold and periodic policies incur same average expected transmission cost, irrespective of the values of E and v. As can be seen from Fig. 2, as expected, the threshold policy has better performance than the periodic policy. Furthermore, as ϵ which represents the error beyond which the estimate is considered incorrect, is increased, the average expected AoIE increases. This is because, as ϵ increases, the *tolerance* for E and v increases, therefore average expected AoIE is higher for a particular average expected transmission cost.

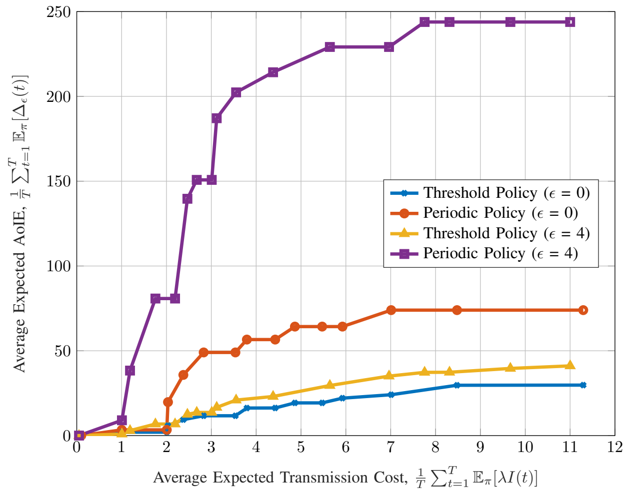

Fig. 2: Variation of the average expected AoIE with the average expected communication cost. The i.i.d. noise, W_t is uniformly distributed over $\{-1, 1\}$, $a = 1$ and the discount factor, $\gamma = 0.99$. The periodic policy transmits periodically after a certain fixed number of slots, K, irrespective of the values of E and v. The value of K is chosen such that the optimal threshold policy and the periodic policy incur same average expected transmission cost.

V. Conclusions

In this work, we considered a source that samples an auto-regressive Markov process and transmits it to a monitor, by incurring a fixed cost. The monitor gains information about the source either by a received sample (when a sample is received) or by estimating the realization of the random process based on the previously received samples (when no samples are received). In this setup, we noted an inadequacy of the classical age of information (AoI) metric, which is, it increases even when the monitor gains an accurate information about the source by estimation. This is because the AoI metric does not account for information content of the received packets. In order to account for this, we adopted a metric referred to as the Age of Incorrect Estimates (AoIE), which is the product of the estimation error, E, and the time elapsed since the latest time the monitor had a sufficiently correct estimate, v. We formulated an optimization problem for deciding when a source must transmit a sample for minimizing the long-term average expected weighted sum of the AoIE and the transmission cost. We proved that the optimal policy for solving the above problem has a threshold structure in E for a fixed v and vice versa.

Appendix

A. Proof of Lemma 2

In order to prove $\mathbb{E}_{\Delta'}[\Delta'|E + e, v, 0] \geq \mathbb{E}_{\Delta''}[\Delta''|E, v, 0]$, we need to compute each of these expected values. Towards this, at time slot t, let us assume that the previous transmission has happened L time slots ago, i.e., at time slot $t - L$. An estimate of $X(t + 1)$ denoted $\hat{X}(t + 1)$ is obtained as follows

$$\hat{X}(t + 1) = a^{L+1} X(t - L). \tag{21}$$

When the current state is (E, v), the square estimation error is given by $E = (a^L X(t - L) - X(t))^2$. Using this and (7), the next state E'' can be written as

$$\begin{aligned} E'' &= \left(\hat{X}(t + 1) - X(t + 1)\right)^2 \\ &= \left(a^{L+1} X(t - L) - (aX(t) + W_{t+1})\right)^2 \\ &= a^2 \left(\pm\sqrt{E} - \frac{W_{t+1}}{a}\right)^2. \end{aligned} \tag{22}$$

ISBN 978-3-903176-37-9 © 2021 IFIP

For a given E, define

$$p_E^+ = \mathbb{P}[\mathcal{E} = +\sqrt{E}|\mathcal{E}^2 = E], \tag{23}$$

and

$$p_E^- = \mathbb{P}[\mathcal{E} = -\sqrt{E}|\mathcal{E}^2 = E], \tag{24}$$

where $\mathcal{E} \triangleq a^L X(t-L) - X(t)$. Conditioned on the event $\mathcal{E} = +\sqrt{E}$, the event $E'' \geq \epsilon$, i.e., $a^2\left(\sqrt{E} - \frac{W_{t+1}}{a}\right)^2 \geq \epsilon$ can be equivalently written as

$$a\sqrt{E} - \epsilon \geq W_{t+1} \text{ or } a\sqrt{E} + \epsilon \leq W_{t+1} \tag{25}$$

Using the above, and the fact that $\mathcal{E}$ is independent of W_{t+1}, we have

$$\begin{aligned}
&\mathbb{P}(E'' \geq \epsilon|\mathcal{E} = +\sqrt{E}, v, 0) = \\
&\mathbb{P}[W_{t+1} \leq a\sqrt{E} - \epsilon] + \mathbb{P}[W_{t+1} \geq a\sqrt{E} + \epsilon] \\
&= 1 - \mathbb{P}[a\sqrt{E} - \epsilon \leq W_{t+1} \leq a\sqrt{E} + \epsilon].
\end{aligned} \tag{26}$$

Similarly,

$$\begin{aligned}
&\mathbb{P}(E'' \geq \epsilon|\mathcal{E} = -\sqrt{E}, v, 0) = \\
&\mathbb{P}[W_{t+1} \leq -a\sqrt{E} - \epsilon] + \mathbb{P}[W_{t+1} \geq -a\sqrt{E} + \epsilon] \\
&= 1 - \mathbb{P}[-a\sqrt{E} - \epsilon \leq W_{t+1} \leq -a\sqrt{E} + \epsilon]
\end{aligned} \tag{27}$$

Using this and the total probability law, we get

$$\begin{aligned}
\mathbb{P}(E'' \geq \epsilon|E, v, 0) &= \mathbb{P}(E'' \geq \epsilon|\mathcal{E} = +\sqrt{E}, v, 0)p_E^+ \\
&\quad + \mathbb{P}(E'' \geq \epsilon|\mathcal{E} = -\sqrt{E}, v, 0)p_E^- \\
&= (1 - \mathbb{P}[a\sqrt{E} - \epsilon \leq W_{t+1} \leq a\sqrt{E} + \epsilon])p_E^+ \\
&\quad + (1 - \mathbb{P}[-a\sqrt{E} - \epsilon \leq W_{t+1} \leq -a\sqrt{E} + \epsilon])p_E^- \\
&\overset{(a)}{=} 1 - \mathbb{P}[a\sqrt{E} - \epsilon \leq W_{t+1} \leq a\sqrt{E} + \epsilon],
\end{aligned} \tag{28}$$

where (a) follows from $\mathbb{P}[a\sqrt{E} - \epsilon \leq W_{t+1} \leq a\sqrt{E} + \epsilon] = \mathbb{P}[-a\sqrt{E} - \epsilon \leq W_{t+1} \leq -a\sqrt{E} + \epsilon]$ for a zero mean symmetric W_{t+1} and $p_E^+ = 1 - p_E^-$.

Recall that $\mu_\epsilon(t)$ is the time at which the last estimate was correct. The expected value of $\Delta'' = Ev$, conditioned on the state, (E, v), and the action, $I = 0$, is given by

$$\begin{aligned}
&\mathbb{E}[\Delta''|E, v, 0] = (t + 1 - \mu_\epsilon(t))\mathbb{E}[E''|E, v, 0] \\
&= (v + 1)\mathbb{E}[E''|E, v, 0, E'' \geq \epsilon]\mathbb{P}(E'' \geq \epsilon|E, v, 0) \\
&\quad + 0 \times \mathbb{E}[E''|E, v, 0, E < \epsilon]\mathbb{P}(E'' \leq \epsilon|E, v, 0), \\
&= (v + 1)\mathbb{E}[E''|E, v, 0, E'' \geq \epsilon]\mathbb{P}(E'' \geq \epsilon|E, v, 0), \\
&\overset{(a)}{=} (v + 1)\frac{\int_0^\infty \mathbb{P}(E'' > t|E, v, 0)dt}{\mathbb{P}(E'' \geq \epsilon|E, v, 0)}\mathbb{P}(E'' \geq \epsilon|E, v, 0), \\
&= (v + 1)\int_0^\infty \mathbb{P}(E'' > t|E, v, 0)dt,
\end{aligned} \tag{29}$$

where (a) follows from

$$\begin{aligned}
\mathbb{E}[E''|E, v, 0, E'' \geq \epsilon] &= \int_0^\infty \mathbb{P}(E'' > t|E, v, 0, E'' \geq \epsilon)dt \\
&= \frac{\int_0^\infty \mathbb{P}(E'' > t|E, v, 0)dt}{\mathbb{P}(E'' \geq \epsilon|E, v, 0)}.
\end{aligned}$$

Similarly, when the current state is $(E + e, v)$, the next state

$$E' = a^2\left(\pm\sqrt{E + e} - \frac{W_{t+1}}{a}\right)^2, \tag{30}$$

and the expected value of $\Delta' = E'v$, conditioned on the state, $(E + e, v)$, and the action, $I = 0$, is given by

$$\mathbb{E}[\Delta'|E + e, v, 0] = (v + 1)\int_0^\infty \mathbb{P}(E' > t|E + e, v, 0)dt, \tag{31}$$

Similarly from (28), we can infer that

$$\begin{aligned}
&\mathbb{P}(E' \geq \epsilon|E + e, v, 0) = \mathbb{P}(E' \geq \epsilon|\mathcal{E} = +\sqrt{E + e}, v, 0)p_{E+e}^+ \\
&\qquad + \mathbb{P}(E' \geq \epsilon|\mathcal{E} = -\sqrt{E + e}, v, 0)p_{E+e}^- \\
&= (1 - \mathbb{P}[a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq a\sqrt{E + e} + \epsilon])p_{E+e}^+ \\
&\quad + (1 - \mathbb{P}[-a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq -a\sqrt{E + e} + \epsilon])p_{E+e}^- \\
&\overset{(a)}{=} 1 - \mathbb{P}[a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq a\sqrt{E + e} + \epsilon],
\end{aligned} \tag{32}$$

where (a) is because $\mathbb{P}[a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq a\sqrt{E + e} + \epsilon] = \mathbb{P}[-a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq -a\sqrt{E + e} + \epsilon]$ for a zero mean symmetric W_{t+1} and $p_{E+e}^+ = 1 - p_{E+e}^-$.

Furthermore, from (8), it follows that $\mathbb{P}[a\sqrt{E + e} - \epsilon \leq W_{t+1} \leq a\sqrt{E + e} + \epsilon] \leq \mathbb{P}[a\sqrt{E} - \epsilon \leq W_{t+1} \leq a\sqrt{E} + \epsilon]$. Now, from (28) and (32), we obtain,

$$\mathbb{P}(E' > t|E + e, v, 0) \geq \mathbb{P}(E'' > t|E, v, 0) \tag{33}$$

Since v has been fixed in (29) and (31), we have, $\mathbb{E}[\Delta'|E + e, v, 0] \geq \mathbb{E}[\Delta''|E, v, 0]$.

B. *Proof of Lemma 3*

We prove the lemma using induction. First, note that $J_{\gamma,0}(E, v) = Ev$ is a monotonically non-decreasing function in E for a fixed v. Now, suppose that $J_{\gamma,t}(E, v)$ is non-decreasing in E for a fixed v. From (33), we can obtain, $\mathbb{P}(E' > x, v + 1|E + e, v, 0) > \mathbb{P}(E'' > x, v + 1|E, v, 0)$, which implies,

$$\mathbb{P}(E' \leq x, v + 1|E + e, v, 0) \leq \mathbb{P}(E'' \leq x, v + 1|E, v, 0).$$

Moreover, $\mathbb{P}(E' \leq x, 1|E + e, v, 1) \leq \mathbb{P}(E'' \leq x, 1|E, v, 1)$. Hence, we have

$$\mathbb{P}(E' \leq x, v'|E + e, v, I) \leq \mathbb{P}(E' \leq x, v'|E, v, I). \tag{34}$$

In words, (34) means that the conditional distribution of the next state given that the current state is $E + e$, first order stochastically dominates that when the current state is E. This implies that if $J_{\gamma,t}(E, v)$ is monotonically non-decreasing function in E for a fixed v, then $\mathbb{E}[J_{\gamma,t}(E', v')|E + e, v, I] \geq \mathbb{E}[J_{\gamma,t}(E', v')|E, v, I]$ [15].

Now since the immediate reward and $\mathbb{E}[J_{\gamma,t}(E', v')|E, v, I]$ are non-decreasing functions of E for fixed v, and since the minimum operator preserves the non-decreasing behavior of functions, $J_{\gamma,t+1}(E, v)$ is non-decreasing in E for a fixed v. Hence, the proof by induction.

ISBN 978-3-903176-37-9 © 2021 IFIP

C. Proof of Lemma 5

We first consider

$$\begin{aligned}\mathbb{E}_{\Delta'}\left[\Delta'+0|E,v+1,0\right]=&\mathbb{E}\left[(v+2)E'+0|E,v+1,0\right]\\=&(v+2)\mathbb{E}[E'|E,v+1,0]\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}_{\Delta''}\left[\Delta''+0|E,v,0\right]=&\mathbb{E}\left[(v+1)E''+0|E,v,0\right]\\=&(v+1)\mathbb{E}[E''|E,v,0].\end{aligned}$$

From (22), we have, $\mathbb{E}[E'|E,v+1,0]=\mathbb{E}[E''|E,v,0]$, as E' and E'' do not depend on v. Hence, $\mathbb{E}_{\Delta'}\left[\Delta'+0|E,v+1,0\right]=\mathbb{E}_{\Delta''}\left[\Delta''+0|E,v,0\right]$.

References

[1] 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.*, vol. 28, no. 5, pp. 2215–2228, 2020.

[2] C. Kam, S. Kompella, and A. Ephremides, "Age of incorrect information for remote estimation of a binary markov source," in *IEEE INFOCOM*, 2020, pp. 1–6.

[3] Y. Chen and A. Ephremides, "Minimizing age of incorrect information for unreliable channel with power constraint," 2021. [Online]. Available: https://arxiv.org/abs/2101.08908

[4] R. D. Yates, Y. Sun, D. R. B. I. au2, S. K. Kaul, E. Modiano, and S. Ulukus, "Age of information: An introduction and survey," 2020. [Online]. Available: https://arxiv.org/abs/2007.08564

[5] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *IEEE SECON*, 2011, pp. 350–358.

[6] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *IEEE INFOCOM*, 2012, pp. 2731–2735.

[7] J. Zhong, R. D. Yates, and E. Soljanin, "Two freshness metrics for local cache refresh," in *IEEE ISIT*, 2018, pp. 1924–1928.

[8] W. Lin, X. Wang, C. Xu, X. Sun, and X. Chen, "Average age of changed information in the internet of things," in *IEEE WCNC*, 2020, pp. 1–6.

[9] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *IEEE ISIT*, 2017, pp. 321–325.

[10] T. Z. Ornee and Y. Sun, "Sampling for remote estimation through queues: Age of information and beyond," in *IEEE WiOPT*, 2019, pp. 1–8.

[11] J. Chakravorty and A. Mahajan, "Fundamental limits of remote estimation of autoregressive markov processes under communication constraints," *IEEE Trans. Autom. Control*, vol. 62, no. 3, pp. 1109–1124, 2017.

[12] Y. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, pp. 1–1, 2019.

[13] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed. USA: John Wiley & Sons, Inc., 1994.

[14] M. Goyal, A. Kumar, and V. Sharma, "Optimal cross-layer scheduling of transmissions over a fading multiaccess channel," *IEEE Trans. on Info. Theory*, vol. 54, no. 8, pp. 3518–3537, 2008.

[15] V. Krishnamurthy, *Partially Observed Markov Decision Processes: Structural Results*. Cambridge University Press, 2016, p. 201–202.

ISBN 978-3-903176-37-9