Minimizing the Age of Information Over an Erasure Channel for Random Packet Arrivals With a Storage Option at the Transmitter

Subhankar Banerjee   Sennur Ulukus   Anthony Ephremides
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
sbanerje@umd.edu   *ulukus@umd.edu*   *etony@umd.edu*

Abstract—We consider a time slotted communication system consisting of a base station (BS) and a user. At each time slot an update packet arrives at the BS with probability p, and the BS successfully transmits the update packet with probability q over an erasure channel. We assume that the BS has a unit size buffer where it can store an update packet upon paying a storage cost c. There is a trade-off between the age of information and the storage cost. We formulate this trade-off as a Markov decision process and find an optimal switching type storage policy.

I. Introduction

We consider a time-slotted communication system, where at each time slot, an update packet arrives at the BS with a geometric distribution and the BS transmits the update to a user over an unreliable channel. We use the recently introduced age of information metric (see [1]–[3] and the references therein) which captures the freshness of information. Most of the works in the literature which consider the minimization of age of information for such a system, i.e., a transmitter and receiver pair (can also have multiple receivers or multiple sources) over an erasure channel with stochastic arrivals of the update packet at the transmitter, has a queuing model present in it to store an update packet while it is being transmitted or to store a new update packet which may arrive while the transmitter is busy with the transmission of an older update packet, see [4]–[16]. These papers consider a diverse set of problems, such as, deriving exact or approximated expressions for average or peak age of information for various transmission policies under different network statistics, or finding the optimal scheduling policies for various network statistics scenarios.

Similar to the aforementioned papers, we too consider that a buffer is present at the transmitter. Note that, as we are dealing with a single source and a single receiver system, and as the packet arrival time and the packet delivery time follow geometric distributions, we can show that a preemptive transmission policy is optimal; a continuous counterpart of this argument is shown in [9]. Thus, a size 1 buffer is sufficient for our problem. However, different than all the above mentioned work, we consider that if the BS stores an update packet to the buffer it has to pay a storage cost c per stored packet.

With the addition of the storage cost, we analyze the trade-off between the storage cost and the age of information, which can be useful in the study of other problems. For example,

consider a limited size buffer transmitter with multiple sources: If at time t, the number of packets that arrive at the transmitter is more than the available buffer size, the transmitter needs to decide which packets to store and which packets to discard. To derive a Whittle index policy [17] for this problem necessitates a mathematical understanding of the trade-off between the storage cost and the age of information.

In this paper, we formulate the minimization of the storage cost plus the age of information as a countable infinite state Markov decision process (MDP) and we show that there exists an optimal switching type policy for this MDP. We then find an optimal switching type storage policy.

II. System Model and Problem Formulation

At each time slot, the BS receives an update packet from the process of interest with probability $0 < p < 1$, and the BS transmits this update packet to the user over an erasure channel with probability of successful transmission $0 < q < 1$. Upon the reception of the update packet, the user transmits an acknowledgement signal to the BS. We assume that this acknowledgement signal reaches the BS error-free and with a one time slot delay, i.e., if the BS transmits an update packet at time slot t, then at time slot $t + 1$ the BS gets to know whether the update packet has been successfully received by the user or not. As at each time slot there is an uncertainty about the arrival of an update packet at the BS, and as the communication channel is erroneous, the BS may want to store the update packet. For this purpose, the BS has a unit-size buffer to store an update packet.

We assume that each time the BS stores an update packet at the buffer it incurs a storage cost c. For the simplicity of this paper, we assume that the buffer drops a stored update packet at the end of one time slot after it is stored. For example, assume that at time slot t the BS receives an update packet, it transmits the update packet to the user and stores the update packet in the buffer with cost c. Also assume that at time slot $t + 1$ the BS does not receive any new update packets, however, it transmits the stored update packet from the buffer, and at the end of time slot $t+1$ the update packet gets dropped from the buffer. With these assumptions, whenever there is a successful transmission of an update packet, the age of the system drops down to either 1 or 2. The more general case,

i.e., when a stored update packet does not get dropped from the buffer and it only gets preempted with a fresher update packet can be an interesting extension of this work; in this general case, the age may drop down to 3, 4, etc as well.

Let the BS employ a storage algorithm π. We denote the action of the BS under the storage policy π at time slot t as $a^\pi(t)$, where $a^\pi(t) = 1$ denotes that the BS stores the update packet at time t and $a^\pi(t) = 0$ denotes otherwise. From the above discussion it is evident that if the instantaneous age of the user at time slot t is $v^\pi(t)$, then at time slot $t + 1$ the instantaneous age is either $v^\pi(t) + 1$ or 1 or 2, where the age $v^\pi(t) + 1$ corresponds to no successful transmission of an update packet, the age 1 corresponds to arrival of a fresh update packet at time t and successful delivery of that packet by the end of time slot t, and finally, the age 2 corresponds to no arrival of a fresh update packet at time t and successful transmission of the stored update packet from the buffer.

A storage policy π is completely defined by the sequence $\{a^\pi(t)\}_{t=1}^\infty$. We define the storage cost corresponding to the policy π at time slot t as $c^\pi(t)$, where $c^\pi(t) = c$ if $a^\pi(t) = 1$ and $c^\pi(t) = 0$ if $a^\pi(t) = 0$. The goal of the transmitter is to find the storage policy π which minimizes the age of information plus the storage cost. We only consider the *causal* policies, i.e., the policies which make a decision based on only the current and the past information. We define the set Π as the set of all causal functions. Formally, the BS considers the following problem

$$\inf_{\pi \in \Pi} \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} v^\pi(t) + c^\pi(t) \right] \qquad (1)$$

III. ALGORITHM AND ANALYSIS

Note that the problem (1) is a countably infinite state MDP. We first define necessary components to solve this MDP.

State: The state of the MDP in (1) is $S = (v, \lambda, b)$. The component v corresponds to the instantaneous age of the user. The component λ corresponds to the availability of a fresh update packet to the BS, $\lambda = 1$ denotes the availability of a fresh update and $\lambda = 0$ denotes otherwise. The component b corresponds to the buffer state, $b = 1$ denotes that an update packet is present in the buffer and $b = 0$ denotes that there is no update packet stored in the buffer. We define $\mathcal{S}$ as the set of all possible states. Note that as the age can be unbounded, the set $\mathcal{S}$ is countably infinite. For a storage policy π, the state of the MDP at time t is $S^\pi(t) = (v^\pi(t), \lambda(t), b^\pi(t))$.

Transition probability: Because of the system model, the considered problem has a time-invariant transition probabilities, i.e., the transition probabilities only depend on the state and the action, and is invariant of the time. Consider two sates in the state space $\mathcal{S}$: $S = (v, \lambda, b)$ and $S' = (v', \lambda', b')$. Under action a, we define the transition probability from state S to state S' as $P_a(S, S')$. First, we consider that $S = (v, 1, b)$, where $b \in \{0, 1\}$, $S' = (1, \lambda', b')$ and $S'' = (v + 1, \lambda', b')$, where $\lambda' \in \{0, 1\}$ and $b' \in \{0, 1\}$, as follows

$$P_1(S, S') = b'(q(\lambda'p + (1 - \lambda')(1 - p))) \qquad (2)$$

$$P_0(S, S') = (1 - b')(q(\lambda'p + (1 - \lambda')(1 - p))) \qquad (3)$$
$$P_1(S, S'') = b'((1 - q)(\lambda'p + (1 - \lambda')(1 - p))) \qquad (4)$$
$$P_0(S, S'') = (1 - b')((1 - q)(\lambda'p + (1 - \lambda')(1 - p))) \qquad (5)$$

Now, consider $S = (v, 0, b)$, $S' = (2, \lambda', 0)$ and $S'' = (v + 1, \lambda', 0)$, where again $\lambda' \in \{0, 1\}$ and $b \in \{0, 1\}$, for any action $a \in \{0, 1\}$,

$$P_a(S, S') = b(q(\lambda'p + (1 - \lambda')(1 - p))) \qquad (6)$$
$$\begin{aligned} P_a(S, S'') = &(1 - b)(\lambda'p + (1 - \lambda')(1 - p)) \\ &+ b((1 - q)(\lambda'p + (1 - \lambda')(1 - p))) \end{aligned} \qquad (7)$$

For any other pair of states, other than the pairs considered in (2)-(7), the transition probability is 0.

Stationary policies: If a policy π is independent of time and depends only on the state of the system, then it is stationary.

Cost: The cost of this MDP is defined as the sum of the storage cost and the age of information. If the state of the system is S and the action is a, where $S \in \mathcal{S}$ and $a \in \{0, 1\}$, then we define the cost as $C(S, a)$. Thus,

$$C(S, a) = ac + \sum_{S' \in \mathcal{S}} v'P_a(S, S') \qquad (8)$$

Similarly, for a policy π, the cost of the system at time slot t is $C(S^\pi(t), a^\pi(t))$.

Now, consider the following optimization problem,

$$\inf_\pi \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} C(S^\pi(t), a^\pi(t)) \right] \qquad (9)$$

As we assume that the age of the user at time slot 1 is 1, it is immediate that (1) and (9) are equivalent.

For an α such that $0 < \alpha < 1$, we define the total discounted cost for a policy π with $S^\pi(1) = S$, where $S \in \mathcal{S}$, as

$$V_\alpha^\pi(S) = \mathbb{E}_\pi \left[\sum_{t=0}^\infty \alpha^t C(S^\pi(t), a^\pi(t)) \Big| S^\pi(1) = S \right] \qquad (10)$$

We define $V_\alpha(S) = \inf_{\pi \in \Pi} V_\alpha^\pi(S)$. If there exists a policy $\pi \in \Pi$, for which $V_\alpha^\pi(S)$ achieves the minimum, we call it an α-optimal policy.

Theorem 1. *There exists an optimal stationary policy π, which minimizes* $\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} v^\pi(t) + c^\pi(t) \right]$.

Proof: Consider a stationary policy $\bar{\pi}$ such that at each time slot the BS stores an update packet with probability $1/2$. For the policy $\bar{\pi}$, all the states in $\mathcal{S}$ can be viewed as states of a Markov chain and we denote this induced Markov chain for the policy $\bar{\pi}$ as M. The transition probability for M, from state S to state S' is $\frac{P_1(S, S') + P_0(S, S')}{2}$. For this Markov chain, we define the cost for state S as $\frac{C(S, 1) + C(S, 0)}{2}$. Note that M is irreducible as all the states are reachable from all the other states and it is aperiodic as well.

Next, we show that M is a positive recurrent Markov chain. From [18, Thm. 1.27], we know that if any one of the states of a Markov chain is positive recurrent then all of the states

are positive recurrent. We show that the state $(1,1,1)$ is a positive recurrent state. We define $\tau_{S,S}$ to be the time needed to reach back to state S from state S for the first time. Formally, $\tau_{S,S} = \inf_{t\geq 1}\{t : S(t) = S | S(0) = S\}$. We can show that $\sum_{t=1}^{\infty} tP(\tau_{(1,1,1),(1,1,1)} = t) < \infty$, where P is the conventional probability measure. Thus, M is a positive recurrent Markov chain. We define $m_{S,S'}$ as the expected total cost till the system reaches state S' for the first time from state S. Now, consider that S is any arbitrary state in $\mathcal{S}$ and S' is $(1,1,1)$. Note that if the BS never schedules an update packet from the buffer, i.e., if there is no fresh update packet, then the BS does not transmit, thus the average cost to go to the state $(1,1,1)$ for the first time increases. Thus, we can get an upper bound on $m_{S,(1,1,1)}$, for all $S \in \mathcal{S}$, and we can show that this upper bound is finite. Thus, $m_{S,(1,1,1)} < \infty$.

From [19, Prop. 5], we claim that $V_\alpha(S)$ is finite, $0 < \alpha < 1$ and $S \in \mathcal{S}$. Again from [19, Prop. 5], we further claim that for every $S \in \mathcal{S}$, there exists a non-negative real number M_S, such that $V_\alpha(S) - V_\alpha((1,1,1)) \leq M_S$, $0 < \alpha < 1$. Further, for all $S \in \mathcal{S}$, there exists an action a, denoted as a_s, such that $\sum_{S'\in\mathcal{S}} P_{a_s}(S,S')M_{S'} < \infty$. As $V_\alpha(S)$ is finite, $S \in \mathcal{S}$, we can say that $0 \leq V_\alpha(1,1,1) < \infty$. Let us define $V_\alpha(1,1,1) = N$, where $N \in \mathbb{R}$. Thus, $V_\alpha(S) - V_\alpha((1,1,1)) \geq -N$, for all $S \in \mathcal{S}$. Thus, from [19, Thm. 1], we have that there exists an optimal stationary policy which minimizes $\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_\pi \left[\sum_{t=1}^{T} v^\pi(t) + c^\pi(t)\right]$. ∎

Switch based policy: We call a policy π a switch based policy, if given that π stores an update packet for state S then π stores an update packet for any other state S' which satisfies $S' \geq S$, where the inequality is component-wise.

Next, we present an important result for $V_\alpha(S)$ [19, Prop. 1].

Lemma 1. *For every state $S \in \mathcal{S}$ and $0 < \alpha < 1$, if $V_\alpha(S)$ is finite then the following relation holds true, $V_\alpha(S) = \min_a\{C(S,a) + \alpha \sum_{S'\in\mathcal{S}} P_a(S,S')V_\alpha(S')\}$.*

Next, we develop a policy iteration method to obtain $V_\alpha(S)$. We define for $n = 0$, $V_{\alpha,0}(S) = 0$, and for $n \geq 1$, $V_{\alpha,n}(S) = \min_a\{C(S,a) + \alpha \sum_{S'\in\mathcal{S}} P_a(S,S')V_{\alpha,n-1}(S')\}$, $S \in \mathcal{S}$; see [19, Prop. 3].

Lemma 2. *For each $0 < \alpha < 1$, we have $V_{\alpha,n}(S) \to V_\alpha(S)$ for every S.*

Recall that $S = (v, \lambda, b)$, where v corresponds to the age, λ corresponds to the availability of a fresh update packet at the BS, and b corresponds to the availability of a stored update packet in the buffer.

Lemma 3. *For a fixed λ and b, $V_\alpha(S)$ is an increasing function of v, i.e., if $S_1 = (v_1, \lambda, b)$ and $S_2 = (v_2, \lambda, b)$ and if $v_2 \geq v_1$, then $V_\alpha(S_2) \geq V_\alpha(S_1)$.*

Proof: We first show that $V_{\alpha,n}(S_2) \geq V_{\alpha,n}(S_1)$, $n \in \mathbb{N}$. Then, the statement of this lemma directly follows from Lemma 2. We show the monotonicity of $V_{\alpha,n}$, by induction. As $V_{\alpha,0}(S) = 0$, $S \in \mathcal{S}$, for $n = 1$, $V_{\alpha,1}(S) = \min_a C(S,a)$.

Note that from (8), it is immediate that $C(S_2,a) \geq C(S_1,a)$, $a \in \{0,1\}$. Thus, $V_{\alpha,1}(S_2) \geq V_{\alpha,1}(S_1)$. Now assume that,

$$V_{\alpha,n-1}(S_2) \geq V_{\alpha,n-1}(S_1) \tag{11}$$

and as before $a \in \{0,1\}$,

$$C(S_2,a) \geq C(S_1,a) \tag{12}$$

combining (11) and (12), we get,

$$\begin{aligned}V_{\alpha,n}(S_2) &= \min_a\{C(S_2,a) + \alpha \sum_{S'\in\mathcal{S}} P_a(S_2,S')V_{\alpha,n-1}(S')\} \\ &\geq \min_a\{C(S_1,a) + \alpha \sum_{S'\in\mathcal{S}} P_a(S_1,S')V_{\alpha,n-1}(S')\} \\ &= V_{\alpha,n}(S_1) \tag{13}\end{aligned}$$

completing the proof. ∎

Let the state be S and the action taken by the BS be a. Then, we define $V_\alpha(S;a) = C(S,a) + \alpha \sum_{S'\in\mathcal{S}} P_a(S,S')V_\alpha(S')$. Note that $V_\alpha(S) = \min_{a\in\{0,1\}} V_\alpha(S;a)$. Next, we state the optimality of a switch type policy. Note that, switching type policy is a stationary policy.

Lemma 4. *There exists a switching type policy which is optimal for $\limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_\pi \left[\sum_{t=1}^{T} C(S^\pi(t), a^\pi(t))\right]$.*

Proof: We first show that for $0 < \alpha < 1$, there exists a switch type policy which is α-optimal. Consider that $\tilde{\pi}$ is an α-optimal policy. Assume that $\tilde{\pi}$ stores an update packet at state $S = (v, 1, 0)$. Thus, $V_\alpha(S;1) - V_\alpha(S;0) \leq 0$, and

$$c + \alpha(1-q)(1-p)(V_\alpha((v+1,0,1)) - V_\alpha((v+1,0,0))) \leq 0 \tag{14}$$

Consider that $S' = (v+1, 1, 0)$. Now, if we can show that $V_\alpha((v+2,0,1)) - V_\alpha((v+2,0,0)) \leq V_\alpha((v+1,0,1)) - V_\alpha((v+1,0,0))$, then it is evident that $V_\alpha(S';1) - V_\alpha(S';0) \leq 0$. Then, by induction, we can argue that if the BS chooses to store an update packet for state S, then the optimal choice for the BS is to store an update packet for state S', where $S' = (v+x, 1, 0)$, where x is a positive integer.

Similar arguments can be made if the current state is $S = (v, 1, 1)$. If the current state is $S = (v, 0, 1)$ or $S = (v, 0, 0)$, then there is no update packet to store. Combining all these, we claim that there exists a switching type policy which is α-optimal. Now, we show that $V_\alpha((v+2,0,1)) - V_\alpha((v+2,0,0)) \leq V_\alpha((v+1,0,1)) - V_\alpha((v+1,0,0))$:

$$\begin{aligned}&V_\alpha((v+1,0,1)) - V_\alpha((v+1,0,0)) \\ &= 2q - q(v+2) - pqV_\alpha((v+2,1,0)) + pqV_\alpha((2,1,0)) \\ &\quad - (1-p)qV_\alpha((v+2,0,0)) + (1-p)qV_\alpha((2,0,0)) \tag{15}\end{aligned}$$

Similarly,

$$\begin{aligned}&V_\alpha((v+2,0,1)) - V_\alpha((v+2,0,0)) \\ &= 2q - q(v+3) - pqV_\alpha((v+3,1,0)) + pqV_\alpha((2,1,0)) \\ &\quad - (1-p)qV_\alpha((v+3,0,0)) + (1-p)qV_\alpha((2,0,0)) \tag{16}\end{aligned}$$

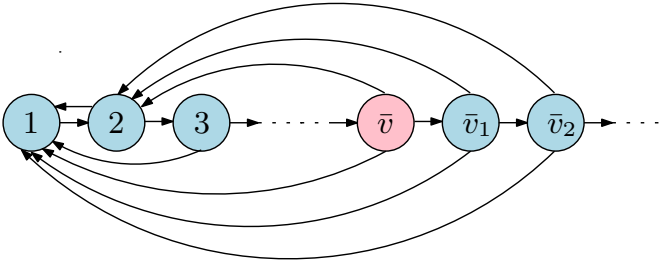From Lemma 3, $V_\alpha((v+3,0,0)) \geq V_\alpha((v+2),0,0)$ and

Fig. 1. Markov chain induced by π_1. In this figure, $\bar{v}_1 = \bar{v}+1$ and $\bar{v}_2 = \bar{v}+2$.

$V_\alpha((v+3,1,0)) \geq V_\alpha((v+2,1,0))$. Thus, from (15) and (16), we have $V_\alpha((v+2,0,1)) - V_\alpha((v+2,0,0)) \leq V_\alpha((v+1,0,1)) - V_\alpha((v+1,0,0))$.

Thus, for $0 < \alpha < 1$, there exists a switch type policy, f_α, which is α-optimal. Now, consider a sequence $\{\alpha_n\}_{n=1}^\infty \subset (0,1)$, such that $\lim_{n\to\infty} \alpha_n = 1$, then from [19, Lemma 1], there exists a subsequence $\{\beta_n\}_{n=1}^\infty$ and a stationary policy f such that $\lim_{n\to\infty} f_{\beta_n}(S) = f(S)$, $S \in \mathcal{S}$. Note that, f is also a switch type policy. Now, from [19, Thm. 1], f is an optimal policy for our problem. $\blacksquare$

According to Lemma 4 there exists a switching type policy which is optimal. Let us call an optimal switching policy as π_1. From Lemma 4, we know that if π_1 stores an update packet to the buffer for state $(v,1,0)$, then it stores an update packet to the buffer for any state $(v_1,1,x)$, where $v_1 \geq v$ and $x \in \{0,1\}$. Thus, we can characterize π_1 only based on the age of the system, i.e., if π_1 stores an available update packet when the age of the system is v, then it stores an available update packet when the age of the system is $v+x$, for $x \geq 0$.

Now, consider that π_1 stores an available update packet when the age of the system is $\bar{v}$ or higher. If we consider a state space consisting of only the age of the system, then the policy π_1 induces an irreducible, aperiodic and positive recurrent Markov chain on that state space, see Fig. 1. Thus, from [18], we know that there exists a unique stationary distribution for this Markov chain. Let us denote the stationary distribution of this Markov chain as $\boldsymbol{h}$, and thus the stationary probability of the age of the system being in state v as h_v.

To find $\boldsymbol{h}$, we assume that the initial distribution of the Markov chain is $\boldsymbol{h}$, and as $\boldsymbol{h}$ is a stationary distribution it is a invariant distribution, i.e., if $\boldsymbol{P}$ is the probability transition matrix then $\boldsymbol{h}\boldsymbol{P}^n = \boldsymbol{h}$, $n \in \mathbb{N}$. First, we find the stationary probability of the age of the system being in state $\bar{v}+i$, $i \geq 1$,

$$h_{\bar{v}+1} = P(v(t) = \bar{v}+1|v(t-1) = \bar{v}) = (1-pq)h_{\bar{v}} \quad (17)$$

For $i > 1$,

$$
\begin{aligned}
h_{\bar{v}+i} =& P(v(t) = \bar{v}+i) \\
=& P(v(t) = \bar{v}+i|v(t-1) = \bar{v}+i-1)h_{\bar{v}+i-1} \\
=& P(v(t) = \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 1) \\
& \cdot P(\lambda(t-1) = 1)h_{\bar{v}+i-1} + P(\lambda(t-1) = 0)h_{\bar{v}+i-1} \\
& \cdot P(v(t) = \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 0)
\end{aligned}
\quad (18)
$$

Now, considering the individual terms in (18) we get,

$$
\begin{aligned}
P(v(t) = \bar{v}+i&|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 1) \\
&= (1-q)
\end{aligned}
\quad (19)
$$

and

$$
\begin{aligned}
P(v(t) &= \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 0) \\
&= P(v(t) = \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 0, \\
&\quad \lambda(t-2) = 0) \cdot P(\lambda(t-2) = 0|v(t-1) = \bar{v}+i-1) \\
&\quad + P(v(t) = \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 0, \\
&\quad \lambda(t-2) = 1) \cdot P(\lambda(t-2) = 1|v(t-1) = \bar{v}+i-1)
\end{aligned}
\quad (20)
$$

further considering the individual terms in (20) we get,

$$
\begin{aligned}
P(\lambda(t-2) &= 1|v(t-1) = \bar{v}+i-1) \\
&= P(v(t-1) = \bar{v}+i-1|\lambda(t-2) = 1, v(t-2) = \\
&\quad \bar{v}+i-2)\frac{P(\lambda(t-2)=1)P(v(t-2) = \bar{v}+i-2)}{h_{\bar{v}+i-1}} \\
&= \frac{p(1-q)h_{\bar{v}+i-2}}{h_{\bar{v}+i-1}}
\end{aligned}
\quad (21)
$$

and

$$
\begin{aligned}
P(\lambda(t-2) &= 0|v(t-1) = \bar{v}+i-1) \\
&= \frac{h_{\bar{v}+i-1} - p(1-q)h_{\bar{v}+i-2}}{h_{\bar{v}+i-1}}
\end{aligned}
\quad (22)
$$

Substituting (21) and (22) in (20), we obtain

$$
\begin{aligned}
P(v(t) &= \bar{v}+i|v(t-1) = \bar{v}+i-1, \lambda(t-1) = 0) \\
&= \frac{h_{\bar{v}+i-1} - pq(1-q)h_{\bar{v}+i-2}}{h_{\bar{v}+i-1}}
\end{aligned}
\quad (23)
$$

Finally, substituting (19) and (23) in (18), we get for $i > 1$,

$$
\begin{aligned}
h_{\bar{v}+i} =& ((1-p) + p(1-q))h_{\bar{v}+i-1} \\
& - (1-p)p(1-q)qh_{\bar{v}+i-2}
\end{aligned}
\quad (24)
$$

Therefore, $h_{\bar{v}+i}$ follows a recursive relation according to (24). Let r_1 and r_2 be the two roots of the following quadratic equation,

$$x^2 - ((1-p) + p(1-q))x + ((1-p)p(1-q)q) = 0 \quad (25)$$

Thus, one solution to the recurrence relation in (24) is,

$$h_{\bar{v}+i} = c_1 r_1^i + c_2 r_2^i \quad (26)$$

where c_1 and c_2 must satisfy the initial conditions, namely,

$$h_{\bar{v}} = c_1 + c_2 \quad (27)$$
$$h_{\bar{v}+1} = c_1 r_1 + c_2 r_2 \quad (28)$$

Solving, (27) and (28) and using (17), we have

$$c_1 = \frac{(r_2 - (1-pq))h_{\bar{v}}}{r_2 - r_1} \quad (29)$$

$$c_2 = \frac{(1-pq)h_{\bar{v}} - r_1}{r_2 - r_1} \quad (30)$$

Inserting these c_1 and c_2 into (26), we obtain, for $i \geq 1$,

$$h_{\bar{v}+i} = \left(\frac{r_2 - (1-pq)}{r_2 - r_1} r_1^i + \frac{(1-pq) - r_1}{r_2 - r_1} r_2^i \right) h_{\bar{v}} \quad (31)$$

Note that (31) is a unique solution to (24) [20, Thm. 2.7]. We define the following quantity which we use in later calculations

$$
\begin{aligned}
\sum_{i=0}^{\infty} h_{\bar{v}+i} &= h_{\bar{v}} \sum_{i=0}^{\infty} \left(\frac{r_2 - 1 + pq}{r_2 - r_1} r_1^i + \frac{1 - pq - r_1}{r_2 - r_1} r_2^i \right) \\
&= \frac{h_{\bar{v}}}{r_2 - r_1} \left(\frac{r_2 - 1 + pq}{1 - r_1} + \frac{1 - pq - r_1}{1 - r_2} \right) \\
&= \frac{h_{\bar{v}}}{pq(1 + (1-p)(1-q))} \quad (32)
\end{aligned}
$$

Now, when $2 \leq j \leq \bar{v}$,

$$h_j = (1 - pq)^{\bar{v}-2} h_2 \quad (33)$$

Note that, the optimal $\bar{v}$ can never be equal to 1, thus $\bar{v} \geq 2$, which is consistent with (31) and (33). Also,

$$h_1 = pq \sum_{i=2}^{\bar{v}-1} h_i + pq \sum_{i=0}^{\infty} h_{\bar{v}+i} \quad (34)$$

Using (32) and (33), we get,

$$h_1 = \left(1 - (1-pq)^{\bar{v}-2} \right) h_2 + \frac{(1-pq)^{\bar{v}-2} h_2}{1 + (1-p)(1-q)} \quad (35)$$

Now, $\sum_{i=1}^{\infty} h_i = 1$, and

$$h_2 = \frac{1}{2 - (1-pq)^{\bar{v}-2} \left(2 - \frac{1}{1+(1-p)(1-q)} \left(\frac{1+pq}{pq} \right) \right)} \quad (36)$$

Combining the results of (31), (33), (35) and (36), we state the following theorem.

Theorem 2. *For a switching type policy π_1 with threshold $\bar{v}$, the stationary distribution of the Markov chain induced by π_1 is given by,*

$$
\begin{aligned}
&h_1 = \left(1 - (1-pq)^{\bar{v}-2} \right) h_2 + \frac{(1-pq)^{\bar{v}-2} h_2}{1 + (1-p)(1-q)} \\
&h_2 = \frac{1}{2 - (1-pq)^{\bar{v}-2} \left(2 - \frac{1}{1+(1-p)(1-q)} \left(\frac{1+pq}{pq} \right) \right)} \\
&h_j = (1-pq)^{\bar{v}-2} h_2, \quad 2 \leq j \leq \bar{v} \\
&h_{\bar{v}+i} = \left(\frac{r_2 - (1-pq)}{r_2 - r_1} r_1^i + \frac{(1-pq) - r_1}{r_2 - r_1} r_2^i \right) h_{\bar{v}}, \quad i \geq 1
\end{aligned}
$$

Now, we find the total cost for the policy π_1. If the age of the system is v, then the corresponding cost of the system is $v + cp$, if $v \geq \bar{v}$ or else the cost is v. Note that, π_1 only stores an update packet if the system receives a fresh update packet with probability p. With a similar argument made for [21, Eqn. (3)], we can say that the cost of the π_1 is $\boldsymbol{h}^T \boldsymbol{c}$, where $\boldsymbol{c}$ is a vector with the jth component being j, if $j < \bar{v}$, otherwise the jth component is $j + cp$. If the age of the system is $\bar{v} + i$, $i \geq 0$, then the corresponding cost for π_1 is $\bar{v} + i + cp$.

Thus,

$$\sum_{i=0}^{\infty} h_{\bar{v}+i}(\bar{v} + i + cp) = dh_{\bar{v}} + \frac{h_{\bar{v}}\bar{v}}{pq(1 + (1-p)(1-q))} \quad (37)$$

where $d = \frac{c}{q(1+(1-p)(1-q))} + \frac{(r_2-1+pq)r_1}{(r_2-r_1)(1-r_1)^2} + \frac{(1-r_1-pq)r_2}{(r_2-r_1)(1-r_2)^2}$. Now, for $\bar{v} > 2$ and the age of the system is v, $2 \leq v \leq \bar{v} - 1$,

$$\sum_{v=2}^{\bar{v}-1} h_v v = h_2 \left(\frac{pq+1}{p^2 q^2} \right) - \frac{h_{\bar{v}}\bar{v}}{pq} - h_{\bar{v}} \left(\frac{pq-1}{p^2 q^2} \right) \quad (38)$$

Thus,

$$
\begin{aligned}
&\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_1} \left[\sum_{t=1}^{T} C(S(t), a^{\pi_1}(t)) \right] \\
&= \sum_{v=1}^{\bar{v}-1} h_v v + \sum_{i=0}^{\infty} h_{\bar{v}+i}(\bar{v} + i + c) \\
&= d_1 h_2 + d_2 h_2 (1-pq)^{\bar{v}-2} \bar{v} + d_3 h_2 (1-pq)^{\bar{v}-2} \quad (39)
\end{aligned}
$$

where, $d_1 = \frac{p^2 q^2 + pq + 1}{p^2 q^2}$, $d_2 = \frac{1}{pq} \left(\frac{1}{1+(1-p)(1-q)} - 1 \right)$ and $d_3 = \left(\frac{d+d(1-p)(1-q)}{1+(1-p)(1-q)} - \frac{1-pq+p^2 q^2}{p^2 q^2} \right)$. Let us first see two limiting cases, i.e., $\bar{v} = 2$ and $\bar{v} = \infty$. For $\bar{v} = 2$,

$$
\begin{aligned}
&\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_1} \left[\sum_{t=1}^{T} C(S(t), a^{\pi_1}(t)) \right] \\
&= \sum_{i=0}^{\infty} h_{2+i}(2 + i + c) + h_1 \\
&= \left(d + \frac{1}{1+(1-p)(1-q)} \left(\frac{pq+2}{pq} \right) \right) h_2 \quad (40)
\end{aligned}
$$

and for $\bar{v} = \infty$,

$$\limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\pi_1} \left[\sum_{t=1}^{T} C(S(t), a^{\pi_1}(t)) \right] = \frac{1 + pq + p^2 q^2}{2p^2 q^2} \quad (41)$$

Now, we find extremal points of the expression of (39), i.e., we find $\bar{v}$ such that the derivative of the function $f(\bar{v}) = d_1 h_2 + d_2 h_2 (1-pq)^{\bar{v}-2} \bar{v} + d_3 h_2 (1-pq)^{\bar{v}-2}$ vanishes at $\bar{v}$. Replacing the expression of h_2 in $f(\bar{v})$, and solving it for 0 we get the following equation of $\bar{v}$,

$$(1-pq)^{\bar{v}-2} = \left(\frac{d_4}{\left(\bar{v} + \frac{1}{\ln(1-pq)} - 1 \right)} + 2 \right) \frac{1}{d_5} \quad (42)$$

where, $d_4 = \frac{\frac{p^2 q^2 + pq + 1}{p^2 q^2} \left(2 - \frac{pq+1}{pq(1+(1-p)(1-q))} \right) + \frac{2}{pq(1+(1-p)(1-q))}}{\frac{1}{pq} \left(\frac{1}{1+(1-p)(1-q)} - 1 \right)} + 2$

and $d_5 = 2 - \frac{1}{1+(1-p)(1-q)} \left(\frac{pq+1}{pq} \right)$.

From (42), the necessary condition for $\bar{v}$ for $f'(\bar{v}) = 0$ is,

$$0 < \frac{d_4}{d_5(\bar{v} + \frac{1}{\ln(1-pq)} - 1)} + \frac{2}{d_5} < 1 \quad (43)$$

Now, if $\frac{d_4}{d_5} > 0$, then the necessary condition for $\bar{v}$ to be an

Algorithm 1 Finding optimal $\bar{v}$

Inputs: h_2, p, q, d, e
Define: $x_1 = \left(d + \frac{1}{1+(1-p)(1-q)}\left(\frac{pq+2}{pq}\right)\right)h_2$, $x_2 = \frac{1+pq+p^2q^2}{2p^2q^2}$, $x_3 = \max\{2, \frac{d_4}{d_5-2} - \frac{1}{\ln(1-pq)} + 1\}$, $x_4 = \max\{2, -\frac{d_4}{2} - \frac{1}{\ln(1-pq)} + 1\}$, $x_5 = 2$, $\mathcal{I}$ is the set of positive integers
if $\frac{c_3}{d_5} > 0$ **then**
    for $x_3 < \bar{v} < x_4$, $\bar{v} \in \mathcal{I}$ **do**
        if $f(\bar{v}) < x_1$ **then**
            $x_1 = f(\bar{v})$
            $x_5 = \bar{v}$
else
    for $x_4 < \bar{v} < x_3$, $\bar{v} \in \mathcal{I}$ **do**
        if $f(\bar{v}) < x_1$ **then**
            $x_1 = f(\bar{v})$
            $x_5 = \bar{v}$
if $x_2 < f(x_5)$ **then**
    Return: $\bar{v} = \infty$
else
    Return: $\bar{v} = x_5$

Fig. 2. Dependence of $\bar{v}$ on varying p, for different q and c.

extremal point is,

$$\max\left\{2, \frac{d_4}{d_5-2} - \frac{1}{\ln(1-pq)} + 1\right\} < \bar{v}$$
$$< \max\left\{2, -\frac{d_4}{2} - \frac{1}{\ln(1-pq)} + 1\right\} \quad (44)$$

and if $\frac{d_4}{d_5} < 0$, then the necessary condition for $\bar{v}$ to be an extremal point is,
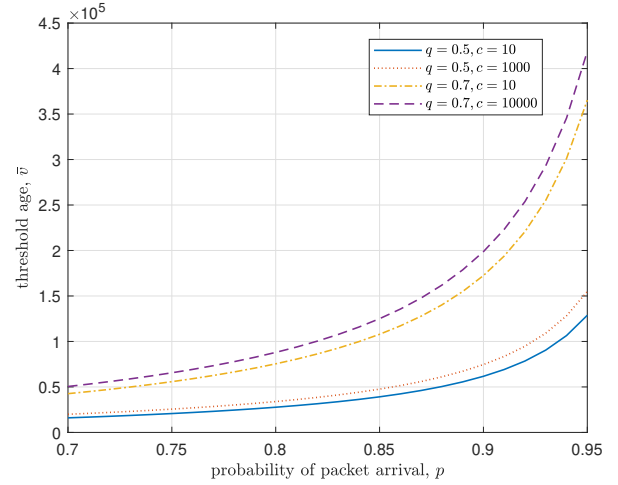
$$\max\left\{2, -\frac{d_4}{2} - \frac{1}{\ln(1-pq)} + 1\right\} < \bar{v}$$
$$< \max\left\{2, \frac{d_4}{d_5-2} - \frac{1}{\ln(1-pq)} + 1\right\} \quad (45)$$

Thus, we evaluate the cost function $f(\bar{v})$, for $\bar{v} = 2$, $\bar{v} = \infty$ and all the integers that satisfy (44) or (45), and choose the $\bar{v}$ which attains the minimum $f(\bar{v})$. This is the optimal $\bar{v}$ as f is a continuous function. This procedure of finding optimal $\bar{v}$ is given in Algorithm 1.

In Fig. 2, we plot the threshold age $\bar{v}$ as a function of p for several fixed values of q and c. Intuitively, if p increases, the probability that the BS has a fresh update packet at each time slot also increases, then paying a storage cost to store an old update packet for lower values of the age becomes sub-optimal, thus the $\bar{v}$ increases with p. Similarly, if q increases, then the probability with which a fresh update packet transmitted by the BS successfully reaches the user also increases, thus storing a stale update packet for lower values of the age becomes sub-optimal, thus the $\bar{v}$ increases with q. Both of these intuitions can be verified with Fig. 2.

REFERENCES

[1] A. Kosta, N. Pappas, and V. Angelakis. Age of information: A new concept, metric, and tool. *Foun. Trend. Netw.*, 12(3):162–259, 2017.

[2] R. D. Yates, Y. Sun, R. Brown, S. K. Kaul, E. Modiano, and S. Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, May 2021.

[3] Y. Sun, I. Kadota, R. Talak, and E. Modiano. Age of information: A new metric for information freshness. *Synthesis Lectures on Communication Networks*, 12(2):1–224, December 2019.

[4] I. Kadota and E. Modiano. Minimizing the age of information in wireless networks with stochastic arrivals. *IEEE Transactions on Mobile Computing*, 20(3):1173–1185, March 2021.

[5] K. Saurav and R. Vaze. Scheduling to minimize age of information with multiple sources. Online arXiv:2204.02953.

[6] M. Moltafet, M. Leinonen, and M. Codreanu. Closed-form expression for the average age of information in a multi-source M/G/1 queueing model. In *IEEE ITW*, August 2019.

[7] K. Saurav and R. Vaze. Minimizing the sum of age of information and transmission cost under stochastic arrival model. In *Infocom*, May 2021.

[8] E. Najm and E. Telatar. Status updates in a multi-stream M/G/1/1 preemptive queue. In *IEEE Infocom*, April 2018.

[9] A. M. Bedewy, Y. Sun, and N. B. Shroff. Minimizing the age of information through queues. *IEEE Transactions on Information Theory*, 65(8):5215–5232, August 2019.

[10] M. Costa, M. Codreanu, and A. Ephremides. On the age of information in status update systems with packet management. *IEEE Transactions on Information Theory*, 62(4):1897–1910, April 2016.

[11] K. Saurav. 3-competitive policy for minimizing age of information in multi-source M/G/1 queuing model. Online at arXiv:2201.03502.

[12] R. D. Yates and S. K. Kaul. The age of information: Real-time status updating by multiple sources. *IEEE Transactions on Information Theory*, 65(3):1807–1827, March 2018.

[13] M. Moltafet, M. Leinonen, and M. Codreanu. On the age of information in multi-source queueing models. *IEEE Transactions on Communications*, 68(8):5003–5017, May 2020.

[14] E. Najm, R. D. Yates, and E. Soljanin. Status updates through M/G/1/1 queues with harq. In *IEEE ISIT*, June 2017.

[15] A. Soysal and S. Ulukus. Age of information in G/G/1/1 systems: Age expressions, bounds, special cases, and optimization. *IEEE Transactions on Information Theory*, 67(11):7477–7489, July 2021.

[16] S. K. Kaul, R. D. Yates, and M. Gruteser. Status updates through queues. In *CISS*, March 2012.

[17] P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25(A):287–298, 1988.

[18] R. Serfozo. *Basics of Applied Stochastic Processes*. Springer Science & Business Media, 2009.

[19] L. I. Sennott. Average cost optimal stationary policies in infinite state Markov decision processes with unbounded costs. *Operations Research*, 37(4):626–633, July 1989.

[20] S. Elaydi. *An Introduction to Difference Equations*, volume 10. Springer-Verlag, New York, 1996.

[21] A. A. Malikopoulos, C. D. Charalambous, and I. Tzortzis. The average cost of Markov chains subject to total variation distance uncertainty. *Systems & Control Letters*, 120:29–35, October 2018.