

To Re-transmit or Not to Re-transmit for Freshness

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 network with a base station (BS) and a user. At each time slot a fresh update packet arrives at the BS with probability $p > 0$. When the BS transmits an update packet for the first time, it goes through with a success probability of q_1 . In all subsequent re-transmissions, the packet goes through with a success probability of q_2 where $q_2 > q_1$, due to the accumulation of observations at the receiver used to decode the packet. When the packet goes through the first time, the age of the user drops to 1, while when the packet goes through in subsequent transmissions, the age of the user drops to the age of the packet since its generation. Thus, when the BS is in the process of re-transmitting an old packet, if it receives a new packet, it has to decide whether to re-transmit the old packet with higher probability of successful transmission but resulting in higher age, or to transmit the new packet which will result in a lower age upon successful reception but this will happen with lower probability. In this paper, we provide an optimal algorithm to solve this problem.

I. INTRODUCTION

We consider a status update system, where a base station (BS) transmits update packets to a user over an erasure channel. Our goal is to minimize the age of information at the user by timely delivery of fresh update packets. Minimizing the age of information in various network settings has been extensively studied in the literature, e.g., [1]–[11]. The problems these papers solve can be broadly categorized into two groups: the first group of problems is to find closed-form expressions for the average or peak age under various network statistics, and the second group of problems is to find optimal scheduling policies for various network settings.

Different than all the aforementioned works, we consider the possibility of re-transmissions with higher probability of successful transmission. We assume that if the BS transmits a fresh update packet, then the user successfully receives it with probability q_1 , and upon the failure of transmission for the first time, the BS can re-transmit the packet with higher probability of successful transmission q_2 . This assumption is valid because the user can use a hybrid automatic repeat request (HARQ) protocol to accumulate previous observations whereby increasing the probability of successful reception.

Minimizing the age of information by introducing coding and HARQ policies to combat the erasure channel for various network settings has been considered extensively in the literature. These works can be broadly classified in two categories: finding the optimal code lengths to minimize the age of information for various network settings [12]–[15], and finding the optimal transmission policies for HARQ techniques under various network settings [16]–[27].

In this paper, we answer the following question: What is the optimal action for the BS when it receives a fresh update packet while an old update packet is being served; should the BS preempt the old packet and transmit the new packet with lower probability of successful transmission q_1 , or should it discard the new packet and continue to transmit the old packet with higher probability of successful transmission q_2 . To find the optimal transmission policy, the above mentioned works [16]–[27] utilize the knowledge of the HARQ policy being used in the system. In this paper, we are agnostic to the underlying HARQ policy employed by the system, which is fundamentally different from these works, and needs a different analysis to find the structure of the optimal policy.¹

In this paper, we show that this problem has an optimal stationary deterministic policy. We find the structure of this optimal policy: Let v_1 denote the age at the user and v_2 denote the age of the old packet at the BS. Let the BS receive a new update packet. Then, we show that, if the optimal action for the BS is to transmit the new packet by preempting the old packet, then for all cases when the age of the user is v_1 and the age of the old packet at the BS is $v_2 + x$ with $x > 0$, the optimal action is also to transmit the new packet by preempting the old one. On the other hand, if the optimal action for the BS is to discard the new packet and keep re-transmitting the old packet, then for all cases when the age at the user is $v_1 + x$ with $x > 0$, the age of the old packet at the BS is v_2 , the optimal action is also to discard the new packet and keep re-transmitting the old one. Finally, leveraging these structural properties, we propose a relative value iteration algorithm to solve the problem optimally.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a slotted wireless communication system, with one BS and one user with Geo/Geo/1/1 queue. The last 1 in the Kendall’s notation denotes that only 1 update packet is kept in the system. At each time slot the BS receives a new update packet with probability $p > 0$. If at time slot t , the queue is empty and the BS receives an update packet, then the BS transmits this update packet with probability of successful transmission of q_1 to the user. If the user does not successfully receive this update packet at time slot t , and if the BS does not receive any new update packet at time slot $t + 1$, then

¹For mathematical tractability, we assume that the successful transmission probability q_2 remains the same for every re-transmission. A more general version of the problem, where the probability of successful transmission depends on the number of re-transmissions, can be an interesting future work.

the BS re-transmits this old update packet with probability of successful transmission q_2 to the user, where $q_2 > q_1$.

The process of re-transmission continues till the BS receives a new update packet or the old update packet which was generated at time t gets received by the user. If the latter case occurs at time slot $t+k$, then the age of the user drops down to k . However, if the former case occurs at time slot $t+k$, then the BS has to decide whether to keep transmitting the old update packet generated at t with probability q_2 or to transmit the new update packet with probability q_1 . At time slot $t+k$, if the BS chooses to transmit the new update packet, the age of the user drops down to 1 with probability q_1 , otherwise, the age of the user drops down to k with probability q_2 .

Consider an algorithm π , which decides at time slot t whether the BS should re-transmit an old update packet or to transmit a newly available update packet, to the user. The action corresponding to the algorithm π , at time slot t is defined as $\pi(t) \in \{0, 1\}$, where 0 means ‘‘re-transmit the old packet’’ and 1 means ‘‘transmit the new packet’’. We call such an algorithm a switching algorithm. We define the set of all possible causal switching algorithms as Π . The age of the user at time t is defined as $v(t) = t - u(t)$, where $u(t)$ is the generation time of the freshest update packet available to the user till time t . We solve the following optimization problem,

$$\inf_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_{\pi}[v(t)] \quad (1)$$

III. OPTIMAL POLICY

In this section, we formulate the problem in (1) as an MDP and solve the MDP by means of value iteration and leveraging the structure of the optimal policy. First, we define the components of the MDP for this problem.

State: The state of this MDP problem is a tuple with three elements (v_1, v_2, b) where v_1 is the age at the user, v_2 is the age of the packet in the BS queue, and b is an indicator variable, where $b = 1$ implies that the BS has a fresh update packet to transmit and 0 implies that it does not. We define \mathcal{S} to be the set of all possible states. Note that the age of the system can possibly be infinite, thus, \mathcal{S} is a countably infinite set. At time t , we denote the state of the system as $s(t)$. For example, if at time t , the last time the user successfully received an update packet was at time t' , and after time t' the BS again received an update packet at time t'' , and from t'' to t the BS kept re-transmitting that packet where $t' < t'' < t$, and also that at time t , the BS has received another fresh update packet, then, $s(t) = (t - t', t - t'', 1)$. If there is no update packet in the queue, then we denote the state as (v_1, ∞, b) .

Action space: We define the action of the BS at time t as $a(t)$, $a(t) \in \{0, 1, 2\}$, where $a(t) = 1$ denotes that at time t , the BS transmits a freshly received update packet, $a(t) = 0$ denotes that the BS re-transmits the old packet in the queue and discards the freshly received update packet (if any), and $a(t) = 2$ denotes that the BS stays idle. Note that, if the state of the system is $(v_1, v_2, 0)$, i.e., that there is an old packet in the queue and no new packet arrives, then the BS always

chooses action 0, i.e., it re-transmits the packet in the queue. Note also that, if the state of the system is $(v_1, \infty, 1)$, i.e., that there is no old packet in the queue, and a new packet arrives, then the BS always chooses action 1, i.e., it transmits the newly arriving packet. Note further that, if the state of the system is $(v_1, \infty, 0)$, i.e., that there is no old packet in the queue and no new packet arrives, then the BS always chooses action 2, i.e., it remains idle.

Transition probabilities: We denote the transition probability from state \hat{s} to state \bar{s} under action a as $P_a(\hat{s}, \bar{s})$. Then, for the following states,

$$\begin{aligned} s &= (v_1, v_2, 0), \quad s_1 = (v_1 + 1, v_2 + 1, 1), \quad s_2 = (v_1 + 1, v_2 + 1, 0) \\ s_3 &= (v_2 + 1, \infty, 0), \quad s_4 = (v_2 + 1, \infty, 1) \end{aligned} \quad (2)$$

we have the transition probabilities,

$$\begin{aligned} P_0(s, s_1) &= p(1 - q_2) \\ P_0(s, s_2) &= (1 - p)(1 - q_2) \\ P_0(s, s_3) &= (1 - p)q_2 \\ P_0(s, s_4) &= pq_2 \end{aligned} \quad (3)$$

Similarly, for the following states,

$$\begin{aligned} s' &= (v_1, v_2, 1), \quad s'_1 = (v_1 + 1, v_2 + 1, 1), \quad s'_2 = (v_1 + 1, v_2 + 1, 0) \\ s'_3 &= (v_2 + 1, \infty, 0), \quad s'_4 = (v_2 + 1, \infty, 1), \quad s'_5 = (v_1 + 1, 1, 0) \\ s'_6 &= (v_1 + 1, 1, 1), \quad s'_7 = (1, \infty, 0), \quad s'_8 = (1, \infty, 1) \end{aligned} \quad (4)$$

we have the transition probabilities,

$$\begin{aligned} P_0(s', s'_1) &= p(1 - q_2) \\ P_0(s', s'_2) &= (1 - p)(1 - q_2) \\ P_0(s', s'_3) &= (1 - p)q_2 \\ P_0(s', s'_4) &= pq_2 \\ P_1(s', s'_5) &= (1 - p)(1 - q_1) \\ P_1(s', s'_6) &= p(1 - q_1) \\ P_1(s', s'_7) &= (1 - p)q_1 \\ P_1(s', s'_8) &= pq_1 \end{aligned} \quad (5)$$

Further, for the following states,

$$\bar{s} = (v_1, \infty, 0), \quad \bar{s}_1 = (v_1 + 1, \infty, 0), \quad \bar{s}_2 = (v_1 + 1, \infty, 1) \quad (6)$$

we have the transition probabilities,

$$\begin{aligned} P_2(\bar{s}, \bar{s}_1) &= (1 - p) \\ P_2(\bar{s}, \bar{s}_2) &= p \end{aligned} \quad (7)$$

Finally, all other transition probabilities are zero.

Cost: If the state of the system is \hat{s} and if the action taken by the BS is a , then we define the cost of the system as $C(\hat{s}, a)$. In this paper, the cost of the system is the age at the user,

$$C(\hat{s}, a) = \sum_{\bar{s} \in \mathcal{S}} P_a(\hat{s}, \bar{s}) \tilde{v} \quad (8)$$

where \tilde{v} is the first component of \bar{s} , i.e., the age at the user. We call the MDP with the above mentioned state space, action

space, transition probabilities and cost as Δ . For a policy π , we define the average cost for Δ starting from state \tilde{s} , as

$$J^\pi(\tilde{s}) = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_\pi [C(s(t), a(t)) | s(0) = \tilde{s}] \quad (9)$$

We define the optimal average cost starting from state \tilde{s} as,

$$J^*(\tilde{s}) = \inf_{\pi \in \Pi} J^\pi(\tilde{s}) \quad (10)$$

Thus, (1) becomes,

$$J^*(\tilde{s}) = \inf_{\pi \in \Pi} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_\pi [C(s(t), a(t)) | s(0) = \tilde{s}] \quad (11)$$

Next, we introduce the corresponding discounted MDP, which we use to solve the original MDP in (11). Consider a policy π and $\alpha \in (0, 1)$, then for state $\tilde{s} \in \mathcal{S}$, we define the discounted cost as $V_{\pi, \alpha}(\tilde{s})$,

$$V_{\pi, \alpha}(\tilde{s}) = \mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \alpha^t C(s(t), a(t)) \middle| s_0 = \tilde{s} \right] \quad (12)$$

We define the optimized $V_{\pi, \alpha}(\tilde{s})$,

$$V_\alpha(\tilde{s}) = \min_{\pi \in \Pi} V_{\pi, \alpha}(\tilde{s}) \quad (13)$$

which is also known as the value function for state \tilde{s} .

Next, we present Lemma 1 and Lemma 2, which are well-known in the MDP literature, see [28], which we will use to find the structure of the optimal policy. We define $V_\alpha(\tilde{s}; a)$ as,

$$V_\alpha(\tilde{s}; a) = C(\tilde{s}, a) + \alpha \sum_{\hat{s} \in \mathcal{S}} P_a(\tilde{s}, \hat{s}) V_\alpha(\hat{s}) \quad (14)$$

Lemma 1. For every $\tilde{s} \in \mathcal{S}$ and for every $\alpha \in (0, 1)$, if $V_\alpha(\tilde{s})$ is finite, then $V_\alpha(\tilde{s})$ satisfies the following equation,

$$V_\alpha(\tilde{s}) = \min_{a \in \{0, 1, 2\}} V_\alpha(\tilde{s}; a) \quad (15)$$

In the MDP literature, (15) is also known as the discounted optimality equation. We define $\forall \tilde{s} \in \mathcal{S}$, $V_{\alpha, 0}(\tilde{s}) = 0$, and

$$V_{\alpha, n}(\tilde{s}; a) = C(\tilde{s}, a) + \alpha \sum_{\hat{s} \in \mathcal{S}} P_a(\tilde{s}, \hat{s}) V_{\alpha, n-1}(\hat{s}) \quad (16)$$

and

$$V_{\alpha, n}(\tilde{s}) = \min_{a \in \{0, 1, 2\}} V_{\alpha, n}(\tilde{s}; a) \quad (17)$$

Lemma 2. For all $\alpha \in (0, 1)$ and $\tilde{s} \in \mathcal{S}$, $V_{\alpha, n}(\tilde{s}) \rightarrow V_\alpha(\tilde{s})$.

We say that f is a stationary deterministic policy for Δ , if it always chooses the same action for a given state \tilde{s} independent of the past decisions, $f(\tilde{s})$ denotes the decision for the state \tilde{s} . In the next theorem, we show the existence of a stationary deterministic policy which is optimal for Δ . From [29, Thm.], we say that under some technical assumptions namely, $V_\alpha(\tilde{s})$ is finite, for all $\tilde{s} \in \mathcal{S}$, (21), (22) and (23), there exists a stationary deterministic policy which is optimal for the MDP Δ . We use [29, Prop. 5] to show these technical assumptions, namely we find a stationary deterministic policy which induces

an ergodic Markov chain and under this policy the expected traveling cost from any state to a fixed state is finite.

Theorem 1. There exists a stationary deterministic policy which minimizes

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_\pi [C(s(t), a(t)) | s(0) = \tilde{s}]. \quad (18)$$

Proof: Consider a stationary policy f , for which the BS only transmits when it receives a fresh update packet, otherwise it stays idle. It is easy to see that the policy f induces a Markov chain M on the state space \mathcal{S} . Note that M is an irreducible and aperiodic Markov chain. Now, we show that M is also positive recurrent. From [30, Thm. 1.27], we know that for an irreducible Markov chain, if one state is positive recurrent, then all of the states are positive recurrent. Now, we show that the state s'_g is positive recurrent. We define $\tau_{\tilde{s}, \tilde{s}}$ to be the time needed to revisit state \tilde{s} starting from the same state \tilde{s} , for the first time, formally $\tau_{\tilde{s}, \tilde{s}} = \inf_{t \geq 1} \{t : s(t) = \tilde{s}, s(0) = \tilde{s}\}$,

$$\mathbb{E}[\tau_{s'_g, s'_g}] = pq_1 + (1 - pq_1) \sum_{i=2}^{\infty} ip^2 q_1 (1 - p^2 q_1)^{i-2} < \infty \quad (19)$$

Now, we define $H_{\hat{s}, \tilde{s}}$ to be the cost that f suffers for visiting state \tilde{s} for the first time starting from state \hat{s} . Then, $H_{\hat{s}, \tilde{s}}$ is a random variable. For example, if $\hat{s} = (2, 1, 1)$ and $\tilde{s} = (1, \infty, 1)$, for policy f , one trajectory is $(2, 1, 1) \rightarrow (3, 1, 1) \rightarrow (1, \infty, 1)$ and the realization of $H_{\hat{s}, \tilde{s}}$ corresponding to this trajectory is $C((3, 1, 1), 1)$. Note that, if $\hat{s} = (\hat{v}_1, \hat{v}_2, b)$, then $C(\hat{s}, 1) \leq \hat{v}_1 + 1$. Now,

$$\mathbb{E}[H_{\hat{s}, s'_g}] \leq (1 - pq_1) p^2 q_1 \sum_{i=1}^{\infty} (1 - p^2 q_1)^{i-1} \sum_{j=1}^i (v_1 + j + 1) < \infty \quad (20)$$

Based on (19) and (20), from [29, Prop. 5], for $0 < \alpha < 1$ and $\tilde{s} \in \mathcal{S}$, $V_\alpha(\tilde{s})$ is finite and for every state \tilde{s} , and that there exists a non-negative real number $M_{\tilde{s}}$, such that

$$V_\alpha(\tilde{s}) - V_\alpha(s'_g) \leq M_{\tilde{s}} \quad (21)$$

Moreover, $\tilde{s} \in \mathcal{S}$, there exists an action $a_{\tilde{s}}$ such that,

$$\sum_{\hat{s} \in \mathcal{S}} P_{a_{\tilde{s}}}(\tilde{s}, \hat{s}) M_{\tilde{s}} < \infty \quad (22)$$

As $V_\alpha(\tilde{s}) < \infty$ for all \tilde{s} , we define $V_\alpha(s'_g) = L$. Thus,

$$V_\alpha(\tilde{s}) - V_\alpha(s'_g) \geq -L, \quad \tilde{s} \in \mathcal{S} \quad (23)$$

Thus, from [29, Thm.], we say that there exists a stationary deterministic policy which is optimal for the MDP Δ . ■

Next, we prove the monotonicity of the value function.

Lemma 3. For any non-negative real valued x, y, v_1 and v_2 ,

$$V_\alpha((v_1 + x, v_2 + y, b)) \geq V_\alpha((v_1, v_2, b)), \quad b \in \{0, 1\} \quad (24)$$

Further,

$$V_\alpha((v_1, \infty, b)) \leq V_\alpha((v_1 + x, \infty, b)) \quad (25)$$

Proof: To be consistent with the state space \mathcal{S} , $v_2 + y \leq v_1 + x$. We prove the first statement of this lemma by induction. Note that, for the first statement we do not consider the action $a = 2$, as the BS always has an update packet to transmit. Recall that, $s' = (v_1, v_2, 1)$ and we define $\tilde{s}_1 = (v_1 + x, v_2 + y, 1)$. Then,

$$V_{\alpha,1}(s'; 1) = C(s', 1) = q_1 + (1 - q_1)(v_1 + 1) \quad (26)$$

$$V_{\alpha,1}(\tilde{s}_1; 1) = C(\tilde{s}_1, 1) = q_1 + (1 - q_1)(v_1 + x + 1) \quad (27)$$

Similarly,

$$V_{\alpha,1}(s'; 0) = C(s', 0) = q_2(v_2 + 1) + (1 - q_2)(v_1 + 1) \quad (28)$$

$$V_{\alpha,1}(\tilde{s}_1; 0) = C(\tilde{s}_1, 0) = q_2(v_2 + y + 1) + (1 - q_2)(v_1 + x + 1) \quad (29)$$

From (26)-(29), we note $V_{\alpha,1}(s'; a) \leq V_{\alpha,1}(\tilde{s}_1; a)$, $a \in \{0, 1\}$. Thus, $V_{\alpha,1}(s') \leq V_{\alpha,1}(\tilde{s}_1)$. Now, assume that the statement is true for $n - 1$, then we show that the statement is true for n ,

$$V_{\alpha,n}(s'; 0) = C(s', 0) + \sum_{\hat{s} \in \mathcal{S}} P_0(s', \hat{s}) V_{\alpha,n-1}(\hat{s}) \quad (30)$$

and similarly,

$$V_{\alpha,n}(\tilde{s}_1; 0) = C(\tilde{s}_1, 0) + \sum_{\tilde{s} \in \mathcal{S}} P_0(\tilde{s}_1, \tilde{s}) V_{\alpha,n-1}(\tilde{s}) \quad (31)$$

Now, we observe that, for a given state, $\hat{s} \in \mathcal{S}$, such that $P_0(s', \hat{s}) > 0$, there exists another state $\tilde{s} \in \mathcal{S}$, such that $\tilde{s} \geq \hat{s}$ and $P_0(s', \hat{s}) = P_0(\tilde{s}_1, \tilde{s})$. Thus, following the induction step we say that,

$$\sum_{\hat{s} \in \mathcal{S}} P_0(s', \hat{s}) V_{\alpha,n-1}(\hat{s}) \leq \sum_{\tilde{s} \in \mathcal{S}} P_0(\tilde{s}_1, \tilde{s}) V_{\alpha,n-1}(\tilde{s}) \quad (32)$$

and from (28) and (29), we know that $C(s', 0) \leq C(\tilde{s}_1, 0)$. Thus, combining these two facts, we have

$$V_{\alpha,n}(s'; 0) \leq V_{\alpha,n}(\tilde{s}_1; 0) \quad (33)$$

Similarly, we have

$$V_{\alpha,n}(s'; 1) \leq V_{\alpha,n}(\tilde{s}_1; 1) \quad (34)$$

Combining these, we obtain

$$V_{\alpha,n}(s') = \min_{a \in \{0,1\}} \{V_{\alpha,n}(s'; a)\} \quad (35)$$

$$\leq \min_{a \in \{0,1\}} \{V_{\alpha,n}(\tilde{s}_1; a)\} \quad (36)$$

$$= V_{\alpha,n}(\tilde{s}_1) \quad (37)$$

Thus, the first part of this lemma follows from Lemma 2. Next, when the state is of the form $(v_1 + x, \infty, b)$, then depending on b , the BS has only one choice, namely, if $b = 1$ the BS always chooses action 1, and if $b = 0$ the BS stays idle, i.e., it chooses action $a = 2$. Now, from Lemma 1 and the first part of this lemma, we have the second part of this lemma. ■

Now, we investigate a structure of the optimal policy.

Lemma 4. *If the optimal policy for the state $s' = (v_1, v_2, 1)$ is to choose action 1, then for $0 \leq x \leq v_1 - v_2$, the optimal policy for the state $(v_1, v_2 + x, 1)$ is also to choose action 1.*

Proof: We first prove this lemma for the α -optimal policy.

From the assumption of this lemma, we have

$$V_{\alpha}(s'; 1) - V_{\alpha}(s'; 0) \leq 0 \quad (38)$$

Expanding the left hand side of (38) and using Lemma 3,

$$V_{\alpha}((v_1, v_2 + x, 1); 1) - V_{\alpha}((v_1, v_2 + x, 1); 0) \leq 0 \quad (39)$$

Thus, for $0 < \alpha < 1$ there exists an α -optimal policy which has the structure given in the statement of the lemma. Now, consider a sequence $\{\alpha_n\}_{n=1}^{\infty} \subset (0, 1)$, such that $\lim_{n \rightarrow \infty} \alpha_n = 1$, then from [29, Lem. 1], there exists a sub-sequence $\{\beta_n\}_{n=1}^{\infty}$ and an optimal stationary policy f such that $\lim_{n \rightarrow \infty} f_{\beta_n}(s) = f(s)$, $s \in \mathcal{S}$. Thus, the stationary policy f also has the structure given in the lemma. ■

The state space \mathcal{S} is countably infinite, to prove Lemma 7 and to apply a value iteration algorithm, we need a finite state space. Thus, we consider an approximating sequence of MDPs, $\{\Delta_N\}_{N=1}^{\infty}$ to approximate Δ . The state space \mathcal{S}_N associated with Δ_N has the states of the form similar to the states for \mathcal{S} , i.e., (v_1, v_2, b) , where $1 \leq v_1 \leq N$, $1 \leq v_2 \leq v_1$, $v_2 = \infty$ and $b \in \{0, 1\}$. Note that, $\mathcal{S}_N \uparrow \mathcal{S}$. The action space for Δ_N is the same as the action space of Δ . To define the transition probabilities for Δ_N , we first define a function $p_{\tilde{s}}(\tilde{s}, \tilde{s}, N)$, note that, $\hat{s} = (\hat{v}_1, \hat{v}_2, \hat{b})$, $\check{s} = (\check{v}_1, \check{v}_2, \check{b})$ and $\tilde{s} = (\tilde{v}_1, \tilde{v}_2, \tilde{b})$.

$$p_{\tilde{s}}(\check{s}, \hat{s}, a, N) = \begin{cases} 1, & \forall \tilde{s} \in \mathcal{S} \setminus \mathcal{S}_N, \hat{v}_1 = \check{v}_1, \check{s} \in \mathcal{S}_N \\ & \hat{v}_2 = \min\{\check{v}_2, \tilde{v}_2\}, \hat{b} = \tilde{b} \\ 0, & \text{otherwise} \end{cases} \quad (40)$$

Now, we define the transition probabilities for Δ_N , i.e., $P_a(\check{s}, \hat{s}; N)$, where $\check{s}, \hat{s} \in \mathcal{S}_N$ and $a \in \{0, 1, 2\}$ as,

$$P_a(\check{s}, \hat{s}; N) = \sum_{\tilde{s} \in \mathcal{S} \setminus \mathcal{S}_N} P_a(\check{s}, \tilde{s}) p_{\tilde{s}}(\check{s}, \tilde{s}, a, N) \quad (41)$$

If $1 \leq \check{v}_1 < N$, $1 \leq \check{v}_2 \leq \check{v}_1$ and $\check{b} \in \{0, 1\}$, then

$$\hat{\mathcal{S}} = \{\hat{s} : P_a(\check{s}, \hat{s}) > 0\} \subset \mathcal{S}_N \quad (42)$$

Thus, for this case, we define $P_a(\check{s}, \hat{s}; N) = P_a(\check{s}, \hat{s})$. From [31], we say that,

$$\lim_{N \rightarrow \infty} P_a(\check{s}, \hat{s}; N) = P_a(\check{s}, \hat{s}) \quad (43)$$

Finally, the cost for the state action pair (\check{s}, a) is the same as the cost for the same state action pair (\check{s}, a) in Δ , $\check{s} \in \mathcal{S}_N$.

Let us define the α -discounted optimal cost and the α -discounted cost corresponding to a policy π for the MDP Δ_N , as $V_{\alpha}^N(\check{s})$ and $V_{\pi, \alpha}^N(\check{s})$ respectively, where $0 < \alpha < 1$ and $\check{s} \in \mathcal{S}_N$. In the next lemma, we show the convergence of a sequence of average cost optimal policies $\{\pi_N\}_{N=1}^{\infty}$ on $\{\Delta_N\}_{N=1}^{\infty}$ to an average cost optimal policy π on Δ . For the definition of convergence of a sequence of policies, see [31]. This lemma will be useful to prove Theorem 2.

To prove this lemma, according to [32], we first show that, for all $N \geq 1$, $0 < \alpha < 1$ and $\check{s} \in \mathcal{S}_N$,

$$V_{\alpha}^N(\check{s}) - V_{\alpha}^N(\check{s}_2) \geq L \quad (44)$$

where $L \in \mathbb{R}$ and $\check{s}_2 = (1, 1, 1)$. Then, we show that $J^*(\check{s}) <$

∞ , $\hat{s} \in \mathcal{S}$. After which, we show that

$$\sum_{\hat{s} \in \mathcal{S}_N} p_{\hat{s}}(\check{s}, \tilde{s}, a, N) V_{\alpha, n}^N(\hat{s}) \leq V_{\alpha, n}^N(\check{s}) \quad (45)$$

where $\check{s} \in \mathcal{S}_N$, $a \in \{0, 1, 2\}$ and $\tilde{s} \notin \mathcal{S}_N$. Finally, we show that there exists a stationary policy f under which, $H_{\hat{s}, \tilde{s}_2}$ and $\tau_{\hat{s}, \tilde{s}_2}$ are finite for all $\hat{s} \in \mathcal{S}$, and

$$\sum_{\hat{s} \in \mathcal{S}_N \setminus \tilde{s}_2} p_{\hat{s}}(\check{s}, \tilde{s}, a, N) H_{\hat{s}, \tilde{s}_2} \leq H_{\check{s}, \tilde{s}_2} \quad (46)$$

where $\tilde{s} \notin \mathcal{S}_N$, $\check{s} \in \mathcal{S}_N$, $H_{\hat{s}, \tilde{s}_2}$ and $\tau_{\hat{s}, \tilde{s}_2}$ are the average traveling cost and the average travel time to the state \tilde{s}_2 from any state \hat{s} , for the first time, under the policy f .

Lemma 5. *If π is a limit point of a sequence of average cost optimal policies, $\{\pi_N\}_{N=1}^{\infty}$ for $\{\Delta_N\}_{N=1}^{\infty}$, then π is an average cost optimal policy for Δ .*

Proof: Let us consider a policy $\tilde{\pi}$, for which the BS remains idle for the entire time horizon. Note that, $V_{\tilde{\pi}, \alpha}^N(\tilde{s}_2)$ is an increasing function of N , thus the limit exists and we denote,

$$\lim_{N \rightarrow \infty} V_{\tilde{\pi}, \alpha}^N(\tilde{s}_2) = V_{\tilde{\pi}, \alpha}^{\infty}(\tilde{s}_2) \quad (47)$$

Thus,

$$V_{\tilde{\pi}, \alpha}^{\infty}(\tilde{s}_2) = \sum_{t=0}^{\infty} \alpha^t C(s(t), 2) \quad (48)$$

$$= \sum_{t=0}^{\infty} \alpha^t (t+2) \leq \frac{1}{1-\alpha} + \frac{1}{(1-\alpha)^2} \quad (49)$$

Thus, $V_{\alpha}^N(\check{s}) - V_{\alpha}^N(\tilde{s}_2) \geq -\left(\frac{1}{1-\alpha} + \frac{1}{(1-\alpha)^2}\right)$, $\check{s} \in \mathcal{S}_N$, $N \geq 1$ and $0 < \alpha < 1$.

Consider the policy f in the proof for Theorem 1, i.e., the BS transmits only if it receives a fresh update packet and remains idle otherwise. Similar to (1) and (11), for the policy f and the initial state \tilde{s}_2 , we say that,

$$J^f(\tilde{s}_2) = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_f[v(t) | s(0) = \tilde{s}_2] \quad (50)$$

Note that, the stochastic process in (50), $\{v(t)\}_{t=1}^{\infty}$ is a renewal process, we define the inter-renewal time as τ_r . We claim that $\mathbb{E}[\tau_r] < \infty$. Thus, from the renewal reward theorem [30], we say that $J^f(\tilde{s}_2) < \infty$. Thus, $J^*(\tilde{s}_2) < \infty$. Note that, f induces an uni-chain on the state space \mathcal{S} , i.e., the induced Markov chain has only one recurrent class plus some transient states. Thus, from [33], we claim that,

$$J^*(\check{s}) = J^*(\tilde{s}_2) < \infty, \quad \check{s} \in \mathcal{S} \quad (51)$$

If $p_{\hat{s}}(\check{s}, \tilde{s}, a, N) > 0$ and $\tilde{v}_2 \neq \infty$, then \hat{s} is of the form $(N, \min\{\tilde{v}_2, \tilde{v}_2\}, \tilde{b})$, note that $\tilde{s} \geq \hat{s}$, thus from the proof of Lemma 3, we say that,

$$\sum_{\hat{s} \in \mathcal{S}_N} p_{\hat{s}}(\check{s}, \tilde{s}, a, N) V_{\alpha, n}^N(\hat{s}) \leq V_{\alpha, n}^N(\check{s}) \quad (52)$$

for $n \geq 1$, $N \geq 1$, $\alpha \in (0, 1)$, $\tilde{s} \notin \mathcal{S}_N$, $\check{s} \in \mathcal{S}_N$ and $a \in$

$\{0, 1, 2\}$. We say that the similar holds true when $\tilde{v}_2 = \infty$.

Consider again the policy f . Under this policy, we can show that $H_{\check{s}, \tilde{s}_2}$ and $\tau_{\check{s}, \tilde{s}_2}$ are finite numbers. From the definition of $p_{\hat{s}}(\check{s}, \tilde{s}, a, N)$ and $H_{\check{s}, \tilde{s}_2}$, we say that,

$$\sum_{\hat{s} \in \mathcal{S}_N \setminus \tilde{s}_2} p_{\hat{s}}(\check{s}, \tilde{s}, a, N) H_{\hat{s}, \tilde{s}_2} \leq H_{\check{s}, \tilde{s}_2} \quad (53)$$

for $n \geq 1$, $N \geq 1$, $\alpha \in (0, 1)$, $\tilde{s} \notin \mathcal{S}_N$ and $\check{s} \in \mathcal{S}_N$. Combining all the above results, the statement of this lemma directly follows from [32]. ■

Note that, $\mathcal{S} = \mathbb{N}^2 \times \{0, 1\}$. Now, we consider another MDP with the state space $\bar{\mathcal{S}} = \mathbb{R}^2 \times \{0, 1\} \cup s_a$. The cost of this new MDP for a state action pair (\bar{s}, a) , is the same as (8), $\bar{s} \in \bar{\mathcal{S}}$. The transition probabilities are defined as, (3), (5) and (7). We call this new MDP as $\bar{\Delta}$. Let us define the discounted optimal cost for $\bar{\Delta}$ as $\bar{V}_{\alpha}(\bar{s})$, $0 < \alpha < 1$, $\bar{s} \in \bar{\mathcal{S}}$. In the next few lemmas, we study further structure of the value function, to get more structural results for the optimal policy.

Lemma 6. *For a state $(v_1, v_2, b) \in \bar{\mathcal{S}}$, the value function $\bar{V}_{\alpha}((v_1, v_2, b))$ is a concave function of v_1 , $v_1 \geq v_2$.*

Lemma 6 can be proved by using mathematical induction.

In a manner similar to constructing $\bar{\mathcal{S}}$ from \mathcal{S} , we construct $\bar{\mathcal{S}}_N$ from \mathcal{S}_N , i.e., $\bar{\mathcal{S}}_N = [0, N]^2 \times \{0, 1\}$. We define a sequence of MDPs $\{\bar{\Delta}_N\}_{N=1}^{\infty}$, from $\{\Delta_N\}_{N=1}^{\infty}$ in a similar manner we defined $\bar{\Delta}$ from Δ . Let us define the discounted optimal cost corresponding to $\bar{\Delta}_N$ as $\bar{V}_{\alpha}^N(\bar{s})$, $\bar{s} \in \bar{\mathcal{S}}_N$.

In the next lemma, we show an interesting structure for the discounted optimal cost for Δ_N , i.e., $V_{\alpha}^N(\check{s})$, $\check{s} \in \mathcal{S}_N$. Specifically, we show (54) and (55). To prove these, we first show that the left derivative of the discounted optimal cost for $\bar{\Delta}_N$ follows (57) and (58), then we integrate both sides of (57) and (58) with proper limits to get (54) and (55).

Lemma 7. *For $y \geq 0$, if $v_1 + y \leq N$ and $v_2 \leq \bar{v}_2 \leq v_1$, $\bar{v}_2 \neq \infty$, then for $N \geq 1$, $b \in \{0, 1\}$,*

$$\begin{aligned} V_{\alpha}^N((v_1 + y, v_2, b)) - V_{\alpha}^N((v_1, v_2, b)) \\ \leq V_{\alpha}^N((v_1 + y, \bar{v}_2, b)) - V_{\alpha}^N((v_1, \bar{v}_2, b)) \end{aligned} \quad (54)$$

Moreover,

$$\begin{aligned} V_{\alpha}^N((v_1 + y, \bar{v}_2, b)) - V_{\alpha}^N((v_1, \bar{v}_2, b)) \\ \leq \left(\frac{1-q_1}{1-q_2}\right) \left(V_{\alpha}^N((v_1 + y, v_2, b)) - V_{\alpha}^N((v_1, v_2, b))\right) \end{aligned} \quad (55)$$

Proof: Note that as v_2 and \bar{v}_2 cannot be infinity, i.e., there is always a packet to transmit for both the states (v_1, v_2, b) and (v_1, \bar{v}_2, b) , we omit the $a = 2$ in the following. We first prove the statement of this lemma for $\bar{V}_{\alpha}^N(\bar{s})$, $\bar{s} \in \bar{\mathcal{S}}_N$. First, note that Lemma 6 is true for $\bar{V}_{\alpha}^N((v_1, v_2, b))$, thus $\bar{V}_{\alpha}^N((v_1, v_2, b))$ is concave in v_1 . Thus, the left derivative of $\bar{V}_{\alpha}^N((v_1, v_2, b))$ with respect to v_1 exists for all $v_1 \in [1, N]$ [34]. Let us define the left derivative of $\bar{V}_{\alpha}^N((v_1, v_2, b))$ with respect to v_1 as

$$\frac{\partial \bar{V}_{\alpha}^N((v_1, v_2, b))^-}{\partial v_1} = \lim_{h \rightarrow 0^+} \frac{\bar{V}_{\alpha}^N((v_1, v_2, b)) - \bar{V}_{\alpha}^N((v_1 - h, v_2, b))}{h} \quad (56)$$

First, we show that for $1 \leq v_1 \leq N$,

$$\frac{\partial \bar{V}_\alpha^N((v_1, v_2, b))^-}{\partial v_1} \leq \frac{\partial \bar{V}_\alpha^N((v_1, \bar{v}_2, b))^-}{\partial v_1} \quad (57)$$

$$\frac{\partial \bar{V}_\alpha^N((v_1, \bar{v}_2, b))^-}{\partial v_1} \leq \left(\frac{1 - q_1}{1 - q_2} \right) \frac{\partial \bar{V}_\alpha^N((v_1, v_2, b))^-}{\partial v_1} \quad (58)$$

We prove both (57) and (58) simultaneously by mathematical induction. We assume that if the optimal action for state $(v_1, v_2, 1)$ is $a = 1$ ($a = 0$), then there exists $\delta > 0$, such that the optimal action for state $(v'_1, v_2, 1)$ is also $a = 1$ ($a = 0$), where $|v'_1 - v_1| \leq \delta$. Note that because of this assumption, the optimal action for state $(v_1, v_2, 1)$ is the same as the optimal action for state $(v_1 - h, v_2, 1)$, for sufficiently small h . Also note that, for state $(v_1, v_2, 0)$ the optimal action is $a = 0$.

Note that Lemma 4 is also true for $\bar{\Delta}_N$, i.e., if the optimal action for state $s' = (v_1, v_2, 1)$ is $a = 1$, then the optimal action for state $\bar{s}_3 = (v_1, \bar{v}_2, 1)$ is also $a = 1$. Note that the above statement is true even for the inductive stages. Now, we assume that $b = 1$ and for the induction step $n = 1$, if the optimal action for state s' is $a = 1$, then from the above, the optimal action at $n = 1$, for state \bar{s}_3 is also 1. Thus,

$$\frac{\partial \bar{V}_{\alpha,1}^N(s')^-}{\partial v_1} = (1 - q_1) = \frac{\partial \bar{V}_{\alpha,1}^N(\bar{s}_3)^-}{\partial v_1} \quad (59)$$

Now, consider the case when the optimal action at $n = 1$ for state \bar{s}_3 is $a = 0$, then again from Lemma 4, the optimal action at $n = 1$ for the state s' is also $a = 0$. Thus,

$$\frac{\partial \bar{V}_{\alpha,1}^N(s')^-}{\partial v_1} = (1 - q_2) = \frac{\partial \bar{V}_{\alpha,1}^N(\bar{s}_3)^-}{\partial v_1} \quad (60)$$

Finally, consider the case when the optimal action at $n = 1$ for state \bar{s}_3 is $a = 1$, however the optimal action at $n = 1$ for state s' is $a = 0$. Then,

$$\frac{\partial \bar{V}_{\alpha,1}^N(s')^-}{\partial v_1} = (1 - q_2) < (1 - q_1) = \frac{\partial \bar{V}_{\alpha,1}^N(\bar{s}_3)^-}{\partial v_1} \quad (61)$$

Combining (59), (60) and (61), we say that,

$$\frac{\partial \bar{V}_{\alpha,1}^N(s')^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,1}^N(\bar{s}_3)^-}{\partial v_1} \quad (62)$$

In a similar fashion, we show the following inequalities,

$$\frac{\partial \bar{V}_{\alpha,1}^N(\bar{s}_3)^-}{\partial v_1} \leq \left(\frac{1 - q_1}{1 - q_2} \right) \frac{\partial \bar{V}_{\alpha,1}^N(s')^-}{\partial v_1} \quad (63)$$

$$\frac{\partial \bar{V}_{\alpha,1}^N((v_1, v_2, 0))^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,1}^N((v_1, \bar{v}_2, 0))^-}{\partial v_1} \quad (64)$$

$$\frac{\partial \bar{V}_{\alpha,1}^N((v_1, \bar{v}_2, 0))^-}{\partial v_1} \leq \left(\frac{1 - q_1}{1 - q_2} \right) \frac{\partial \bar{V}_{\alpha,1}^N((v_1, v_2, 0))^-}{\partial v_1} \quad (65)$$

Now, assume that all the inequalities are true for the inductive stage $n - 1$. We first prove the inequalities for the induction stage n , for $v_1 \leq N - 1$. First, we consider the case when the optimal action for both states s' and \bar{s}_3 at the induction stage n is $a = 0$. Then, for this case we write,

$$\bar{V}_{\alpha,n}^N(s') = q_2(v_2 + 1) + (1 - q_2)(v_1 + 1)$$

$$\begin{aligned} & + pq_2 \bar{V}_{\alpha,n-1}^N((v_2 + 1, \infty, 1)) \\ & + (1 - p)q_2 \bar{V}_{\alpha,n-1}^N((v_2 + 1, \infty, 0)) \\ & + (1 - p)(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, 0)) \\ & + p(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, 1)) \end{aligned} \quad (66)$$

Similarly,

$$\begin{aligned} \bar{V}_{\alpha,n}^N(\bar{s}_3) & = q_2(\bar{v}_2 + 1) + (1 - q_2)(v_1 + 1) \\ & + pq_2 \bar{V}_{\alpha,n-1}^N((\bar{v}_2 + 1, \infty, 1)) \\ & + (1 - p)q_2 \bar{V}_{\alpha,n-1}^N((\bar{v}_2 + 1, \infty, 0)) \\ & + (1 - p)(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, \bar{v}_2 + 1, 0)) \\ & + p(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, \bar{v}_2 + 1, 1)) \end{aligned} \quad (67)$$

Now, from the assumption of the induction stage $n - 1$,

$$\frac{\partial \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, 1))^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,n-1}^N((v_1 + 1, \bar{v}_2 + 1, 1))^-}{\partial v_1} \quad (68)$$

Thus, from (66), (67) and (68), we say that for the first case,

$$\frac{\partial \bar{V}_{\alpha,n}^N(s')^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} \quad (69)$$

Similarly, we show that for the first case,

$$\frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} \leq \left(\frac{1 - q_1}{1 - q_2} \right) \frac{\partial \bar{V}_{\alpha,n}^N(s')^-}{\partial v_1} \quad (70)$$

Further, we show the inequalities when the optimal actions for both states s' , \bar{s}_3 is $a = 1$, at the induction stage n .

Now, consider the case when the optimal action for s' is $a = 0$, however, the optimal action for \bar{s}_3 is $a = 1$. Then,

$$\begin{aligned} \bar{V}_{\alpha,n}^N(s') & = q_2(v_2 + 1) + (1 - q_2)(v_1 + 1) \\ & + pq_2 \bar{V}_{\alpha,n-1}^N((v_2 + 1, \infty, 1)) \\ & + (1 - p)q_2 \bar{V}_{\alpha,n-1}^N((v_2 + 1, \infty, 0)) \\ & + (1 - p)(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, 0)) \\ & + p(1 - q_2) \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, 1)) \end{aligned} \quad (71)$$

Similarly,

$$\begin{aligned} \bar{V}_{\alpha,n}^N(\bar{s}_3) & = q_1 + (1 - q_1)(v_1 + 1) + pq_1 \bar{V}_{\alpha,n-1}^N((1, \infty, 1)) \\ & + (1 - p)q_1 \bar{V}_{\alpha,n-1}^N((1, \infty, 0)) \\ & + p(1 - q_1) \bar{V}_{\alpha,n-1}^N((v_1 + 1, 1, 1)) \\ & + (1 - p)(1 - q_1) \bar{V}_{\alpha,n-1}^N((v_1 + 1, 1, 0)) \end{aligned} \quad (72)$$

Now, from the assumption on the induction step $n - 1$,

$$\frac{\partial \bar{V}_{\alpha,n-1}^N((v_1 + 1, v_2 + 1, b))^-}{\partial v_1} \leq \left(\frac{1 - q_1}{1 - q_2} \right) \frac{\partial \bar{V}_{\alpha,n-1}^N((v_1 + 1, 1, b))^-}{\partial v_1} \quad (73)$$

Thus, from (71) and (72), we say that for this case at the induction stage n ,

$$\frac{\partial \bar{V}_{\alpha,n}^N(s')^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} \quad (74)$$

Now, combining all the above mentioned three cases, we

say that (74) is always true. By similar techniques, we show the the proofs of the following inequalities,

$$\frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} \leq \left(\frac{1-q_1}{1-q_2} \right) \frac{\partial \bar{V}_{\alpha,n}^N(s')^-}{\partial v_1} \quad (75)$$

$$\frac{\partial \bar{V}_{\alpha,n}^N((v_1, v_2, 0))^-}{\partial v_1} \leq \frac{\partial \bar{V}_{\alpha,n}^N((v_1, \bar{v}_2, 0))^-}{\partial v_1} \quad (76)$$

$$\frac{\partial \bar{V}_{\alpha,n}^N((v_1, \bar{v}_2, 0))^-}{\partial v_1} \leq \left(\frac{1-q_1}{1-q_2} \right) \frac{\partial \bar{V}_{\alpha,n}^N((v_1, v_2, 0))^-}{\partial v_1} \quad (77)$$

For the case $(N-1) < v_1 \leq N$, we can show similar inequalities as in (74)-(77). Thus, for $n \geq 1$, $N \geq 1$, $1 \leq v_1 \leq N$, (74)-(77) holds true. Now, we want to find the limiting values for (74)-(77). Next, we want to show the following,

$$\lim_{n \rightarrow \infty} \frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} = \frac{\partial \bar{V}_{\alpha}^N(\bar{s}_3)^-}{\partial v_1} \quad (78)$$

and the rest of the inequalities follow similarly. Note that,

$$\bar{V}_{\alpha,n}^N(\check{s}) \leq \frac{N}{1-\alpha}, \quad \check{s} \in \bar{\mathcal{S}}_N \quad (79)$$

Thus, from [35, Prop. 1], we say that,

$$|\bar{V}_{\alpha}^N(\check{s}) - \bar{V}_{\alpha,n}^N(\check{s})| \leq \frac{\alpha^{n+1}N}{1-\alpha}, \quad n \geq 1, \check{s} \in \bar{\mathcal{S}}_N \quad (80)$$

Thus, the sequence $\{\bar{V}_{\alpha,n}^N(\check{s})\}_{n=1}^{\infty}$ is uniformly convergent on $\bar{\mathcal{S}}_N$. Therefore,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{\partial \bar{V}_{\alpha,n}^N(\bar{s}_3)^-}{\partial v_1} \\ &= \lim_{n \rightarrow \infty} \lim_{h \rightarrow 0^+} \frac{\bar{V}_{\alpha,n}^N((v_1, \bar{v}_2, 1)) - \bar{V}_{\alpha,n}^N((v_1-h, \bar{v}_2, 1))}{h} \quad (81) \end{aligned}$$

$$= \lim_{h \rightarrow 0^+} \lim_{n \rightarrow \infty} \frac{\bar{V}_{\alpha,n}^N((v_1, \bar{v}_2, 1)) - \bar{V}_{\alpha,n}^N((v_1-h, \bar{v}_2, 1))}{h} \quad (82)$$

$$= \frac{\partial \bar{V}_{\alpha}^N(\bar{s}_3)^-}{\partial v_1} \quad (83)$$

where (82) follows from the Moore-Osgood theorem.

For the next part of the proof, we write $\bar{V}_{\alpha}^N((v_1, v_2, 1)) = g(v_1)$, only a function of the age of the user, v_1 . Let us define, $\epsilon_1 > 0$ and $\epsilon_2 > 0$, which are small enough such that $[1 + \epsilon_1, N - \epsilon_2]$ is a non-empty interval. From Lemma 6, we know that $g(v_1)$ is a concave function of v_1 . Thus, $\frac{\partial g(v_1)^-}{\partial v_1}$ is monotonically decreasing and g is continuous on $(1, N)$ [34]. Combining these two statements we say that $\frac{\partial g(v_1)^-}{\partial v_1}$ is finite on $[1 + \epsilon_1, N - \epsilon_2]$, thus $\frac{\partial g(v_1)^-}{\partial v_1}$ is bounded by $\frac{\partial g(1+\epsilon_1)^-}{\partial v_1}$. Consider $x_1, x_2 \in [1 + \epsilon_1, N - \epsilon_2]$ such that $x_1 < x_2$. Take $h > 0$, small enough such that $1 < x_1 - h$ and $x_1 < x_2 - h$. Now, from the convexity of $g(v_1)$,

$$\frac{g(x_1) - g(x_1-h)}{h} \geq \frac{g(x_2) - g(x_1)}{x_2 - x_1} \geq \frac{g(x_2) - g(x_2-h)}{h} \quad (84)$$

thus,

$$\left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_1} \geq \frac{g(x_2) - g(x_1)}{x_2 - x_1} \geq \left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_2} \quad (85)$$

Consider a partition P on $[1 + \epsilon_1, N - \epsilon_2]$, $x_0 = 1 + \epsilon_1 < x_1 < x_2 < \dots < x_n = N - \epsilon_2$. Then, from (85),

$$\begin{aligned} & \sum_{i=1}^n \left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_{i-1}} (x_i - x_{i-1}) \geq \sum_{i=1}^n g(x_i) - g(x_{i-1}) \\ & \geq \sum_{i=1}^n \left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_i} (x_i - x_{i-1}) \quad (86) \end{aligned}$$

Thus,

$$\begin{aligned} U \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right) & \geq g(N - \epsilon_2) - g(1 + \epsilon_1) \\ & \geq L \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right) \quad (87) \end{aligned}$$

where $U \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right)$ and $L \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right)$ are the upper and the lower Riemann sums for the partition P . Again, consider the same partition and assume that $\max_{1 \leq i \leq n} (x_i - x_{i-1}) \leq \epsilon_3$. As $\frac{\partial g(v_1)^-}{\partial v_1}$ is monotone, we say that,

$$\begin{aligned} & \left| U \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right) - L \left(\frac{\partial g(v_1)^-}{\partial v_1}, P \right) \right| \\ & \leq \left| \sum_{i=1}^n \left(\left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_{i-1}} - \left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{x_i} \right) (x_i - x_{i-1}) \right| \quad (88) \end{aligned}$$

$$\leq \epsilon_3 \left| \left(\left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{N-\epsilon_2} - \left. \frac{\partial g(v_1)^-}{\partial v_1} \right|_{1+\epsilon_1} \right) \right| \quad (89)$$

As $\frac{\partial g(v_1)^-}{\partial v_1}$ is also bounded on $[1 + \epsilon_1, N - \epsilon_2]$, from (89), there exists a sequence of partitions $\{P_n\}_{n=1}^{\infty}$, such that,

$$\lim_{n \rightarrow \infty} \left| L \left(\frac{\partial g(v_1)^-}{\partial v_1}, P_n \right) - U \left(\frac{\partial g(v_1)^-}{\partial v_1}, P_n \right) \right| = 0 \quad (90)$$

Now, from (87), we say that,

$$\lim_{n \rightarrow \infty} U \left(\frac{\partial g(v_1)^-}{\partial v_1}, P_n \right) = \lim_{n \rightarrow \infty} L \left(\frac{\partial g(v_1)^-}{\partial v_1}, P_n \right) \quad (91)$$

$$= \int_{1+\epsilon_1}^{N-\epsilon_2} \frac{\partial g(v_1)^-}{\partial v_1} dv_1 \quad (92)$$

$$= g(N - \epsilon_2) - g(1 + \epsilon_1) \quad (93)$$

Now, as g is a continuous function of v_1 , we first take $\epsilon_1 \rightarrow 0$, then take $\epsilon_2 \rightarrow 0$, and finally, we get,

$$\int_1^N \frac{\partial g(v_1)^-}{\partial v_1} dv_1 = g(N) - g(1) \quad (94)$$

From (74), (83) and with similar argument as above,

$$\int_{v_1}^{v_1+y} \frac{\partial \bar{V}_{\alpha}^N(x, v_2, 1)^-}{\partial v_1} dx \leq \int_{v_1}^{v_1+y} \frac{\partial \bar{V}_{\alpha}^N(x, \bar{v}_2, 1)^-}{\partial v_1} dx \quad (95)$$

Thus,

$$\begin{aligned} & \bar{V}_{\alpha}^N((v_1 + y, v_2, b)) - \bar{V}_{\alpha}^N((v_1, v_2, b)) \\ & \leq \bar{V}_{\alpha}^N((v_1 + y, \bar{v}_2, b)) - \bar{V}_{\alpha}^N((v_1, \bar{v}_2, b)) \quad (96) \end{aligned}$$

In a similar fashion, we show all the inequalities considered

in the statement of this lemma for $\bar{V}_\alpha^N(\cdot)$. Now, if $v_1, v_1 + y, v_2, \bar{v}_2 \in \mathbb{N}$, $\bar{V}_\alpha^N(\cdot)$ and $V_\alpha^N(\cdot)$ are the same, this concludes the proof of this lemma. ■

Next, we study another structure of the optimal policy.

Theorem 2. *If $v_1 \geq \frac{q_2 v_2}{q_2 - q_1}$, and if the optimal action for the state $(v_1, v_2, 1)$ is to choose action 0, then for any non-negative x , the optimal action for the state $(v_1 + x, v_2, 1)$ is also to choose action 0.*

Proof: Due to space restrictions, we only provide a sketch of the proof here: For $N \geq 1$, we use induction, Lemma 7 and arguments similar to those in the proof for Lemma 4 to prove the statement of this theorem for Δ_N . Then, we use Lemma 5 to conclude the proof of this theorem. ■

For an MDP, if all the stationary policies induce a uni-chain, then we say that the MDP is uni-chain. In the next theorem we show that the Δ is an uni-chain MDP.

Theorem 3. *Δ is an uni-chain MDP.*

Proof: We only provide a sketch of the proof here: Note that, any policy which transmits when a fresh update packet is available can reach the state $(1, \infty, 1)$ and from this state every other state is reachable. Thus, Δ is a uni-chain MDP. ■

As the MDP Δ is a uni-chain, next, we use the well-known relative value iteration algorithm (RVIA) [28] to find the optimal stationary deterministic algorithm for (11). For $N \geq 1$, RVIA for Δ_N and for all the states $\hat{s} \in \mathcal{S}_N$ is,

$$V_n^N(\hat{s}) = \min_{a \in \{0,1,2\}} \left\{ C(\hat{s}, a) + \sum_{\tilde{s} \in \mathcal{N}} P_a(\hat{s}, \tilde{s}) V_{n-1}^N(\tilde{s}) - V_{n-1}^N((1, 1, 1)) \right\} \quad (97)$$

where $V_n^N(\hat{s})$ is the value function corresponding to the average cost optimality criterion for Δ_N . We initialize $V_0^N(s) = 0$, for all $s \in \mathcal{S}_N$. Now, from Lemma 5, for large N , we obtain the optimal stationary deterministic policy for (11). We use the structural properties of the optimal policy, i.e., Lemma 4 and Theorem 2 to reduce the complexity of RVIA in (97).

REFERENCES

- [1] 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.
- [2] 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.
- [3] E. Najm and E. Telatar. Status updates in a multi-stream M/G/1 preemptive queue. In *IEEE Infocom*, April 2018.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] A. Soysal and S. Ulukus. Age of information in G/G/1 systems: Age expressions, bounds, special cases, and optimization. *IEEE Transactions on Information Theory*, 67(11):7477–7489, July 2021.
- [9] S. K. Kaul, R. D. Yates, and M. Gruteser. Status updates through queues. In *CISS*, March 2012.
- [10] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor. Online timely status updates with erasures for energy harvesting sensors. In *Allerton Conference on Communication, Control, and Computing*, October 2018.
- [11] A. Arafa, J. Yang, S. Ulukus, and H. V. Poor. Timely status updating over erasure channels using an energy harvesting sensor: Single and multiple sources. *IEEE Transactions on Green Communications and Networking*, 6(1):6–19, March 2021.
- [12] E. Najm, E. Telatar, and R. Nasser. Optimal age over erasure channels. *IEEE Transactions on Information Theory*, 68(9):5901–5922, September 2022.
- [13] H. Sac, T. Bacinoglu, E. Uysal-Biyikoglu, and G. Durisi. Age-optimal channel coding blocklength for an M/G/1 queue with harq. In *IEEE SPAWC*, July 2018.
- [14] X. Chen and S. S. Bidokhti. Benefits of coding on age of information in broadcast networks. In *IEEE ITW*, August 2019.
- [15] J. You, S. Wu, Y. Deng, J. Jiao, and Q. Zhang. An age optimized hybrid ARQ scheme for polar codes via Gaussian approximation. *IEEE Wireless Communications Letters*, 10(10):2235–2239, October 2021.
- [16] E. Najm, R. D. Yates, and E. Soljanin. Status updates through M/G/1 queues with harq. In *IEEE ISIT*, June 2017.
- [17] R. D. Yates, E. Najm, E. Soljanin, and J. Zhong. Timely updates over an erasure channel. In *IEEE ISIT*, June 2017.
- [18] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. On timely channel coding with hybrid ARQ. In *IEEE Globecom*, December 2019.
- [19] E. T. Ceran, D. Gündüz, and A. György. A reinforcement learning approach to age of information in multi-user networks with harq. *IEEE Journal on Selected Areas in Communications*, 39(5):1412–1426, May 2021.
- [20] E. T. Ceran, D. Gündüz, and A. György. Average age of information with hybrid ARQ under a resource constraint. *IEEE Transactions on Wireless Communications*, 18(3):1900–1913, March 2019.
- [21] S. Feng and J. Yang. Age-optimal transmission of rateless codes in an erasure channel. In *IEEE ICC*, May 2019.
- [22] S. Farazi, A. G. Klein, and D. R. Brown. Average age of information in update systems with active sources and packet delivery errors. *IEEE Wireless Communications Letters*, 9(8):1164–1168, August 2020.
- [23] Y. Wang, S. Wu, J. Jiao, W. Wu, Y. Wang, and Q. Zhang. Age-optimal transmission policy with harq for freshness-critical vehicular status updates in space-air-ground-integrated networks. *IEEE Internet of Things Journal*, 9(8):5719–5729, April 2020.
- [24] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Timely estimation using coded quantized samples. In *IEEE ISIT*, June 2020.
- [25] A. Arafa, K. Banawan, K. G. Seddik, and H. V. Poor. Sample, quantize, and encode: Timely estimation over noisy channels. *IEEE Transactions on Communications*, 69(10):6485–6499, October 2021.
- [26] A. Li, S. Wu, J. Jiao, N. Zhang, and Q. Zhang. Age of information with hybrid-ARQ: A unified explicit result. *IEEE Transactions on Communications*, 70(12):7899–7914, December 2022.
- [27] A. Baknina and S. Ulukus. Coded status updates in an energy harvesting erasure channel. In *CISS*, March 2018.
- [28] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, 2014.
- [29] 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.
- [30] R. Serfozo. *Basics of Applied Stochastic Processes*. Springer Science & Business Media, 2009.
- [31] L. I. Sennott. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley & Sons, 2009.
- [32] L. I. Sennott. On computing average cost optimal policies with application to routing to parallel queues. *Mathematical Methods of Operations Research*, 45:45–62, 1997.
- [33] D. Bertsekas. *Dynamic Programming and Optimal Control: Volume II*. Athena Scientific, 2012.
- [34] D. Bertsekas. *Convex Optimization Theory: Volume I*. Athena Scientific, 2009.
- [35] S. Ross. *Introduction to Stochastic Dynamic Programming*. Academic Press, 2014.