ABSTRACT

Title of Thesis: THE FUNDAMENTAL NATURE OF AGE OF INCORRECT INFORMATION

Yutao Chen
Master of Science, 2020

Thesis Directed by: Professor Anthony Ephremides
Department of Electrical and Computer Engineering

Age of Incorrect Information (AoII) is a newly introduced metric that captures not only the freshness of information but also the information content of the transmitted packets and the knowledge at the monitor. It overcomes the shortcomings of Age of Information (AoI) in many applications that involve the problem of remotely estimating an event in real-time. However, the fundamental nature of AoII has been elusive so far. This thesis considers a system in which a transmitter sends updates about a Markovian source to a remote monitor through an unreliable channel. By leveraging the notion of Markov Decision Process (MDP), it is shown that a simple "always update" policy minimizes the AoII. The performances of "always update" policy as well as a more general transmission policy - "threshold update" policy are analyzed in this thesis. Lastly, numerical results that highlight the effects of the parameters on the performances of these two transmission policies are provided.

THE FUNDAMENTAL NATURE OF AGE OF INCORRECT INFORMATION

by

Yutao Chen

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2020

Advisory Committee:
Professor Anthony Ephremides, Chair/Advisor
Professor Sennur Ulukus
Associate Professor Behtash Babadi

Acknowledgments

First and foremost, I want to thank my adviser Dr. Anthony Ephremides for his guidance and support throughout this research work. I would like to thank Ali Maatouk for his important feedbacks and inspiring discussions. I would also like to thank Dr. Ulukus and Dr. Babadi for serving on my thesis committee.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Applications, such as autonomous vehicles, control systems, and unmanned aerial vehicles (UAVs), rely heavily on the exchanging of time-sensitive information. In these applications, the freshness of information is critical. Conventional metrics such as throughput and delay are not always optimal when considering the freshness of information. The Age of Information (AoI) introduced in [1] offers a new way to quantify the freshness of information. Let $U(t)$ be the generation time of the last received packet. AoI is a function defined by $\Delta(t) = t - U(t)$.

Recently, research on AoI has been growing fast. AoI in queueing networks is of great interest. In [1] and [2], first-come-first-served (FCFS) queues and last-come-first-served (LCFS) queues are examined. AoI is analyzed in [3] when the source randomly generates status update messages. In [4], AoI in a system where the source node can manage the arriving samples is studies. At the same time, minimizing AoI in wireless broadcast networks is another fundamental problem and attracts the attention of researchers. [5] studied an age minimization problem over a wireless broadcast network with many users. [6] considered a wireless broadcast network with a base station sending information to several clients. The authors in [7] analyzed the performance of the whittle's index policy in a system where a central entity allows only part of the users to transmit simultaneously. Although AoI has been successful in many applications, it shows weakness in many other scenarios.

In the case where an event needs to be monitored, AoI will increase whether the information at the monitor is correct or not. Since we want to keep the information at the monitor as fresh as possible, we have no reason to think the information is obsolete if it is correct. At the same time, AoI will increase at the same pace regardless of the knowledge at the monitor. This will make AoI performs poorly in some applications such as UAVs. If the controller keeps receiving inaccurate information or the gap between the information at the controller and the correct information keeps widening, the UAV will get further and further from the correct track. In both cases, a higher penalty should be paid.

To overcome the shortcomings of AoI, a new metric Age of Incorrect Information (AoII) is introduced in [8]. Let X_t be the true state of the event we want to monitor and $\hat{X}_t$ be the estimated state at the monitor, AoII is defined as $\Delta(t) = f(t) \times g(X_t, \hat{X}_t)$. $f(t)$ can be any increasing time function and $g(X_t, \hat{X}_t)$ can be any function that reflects the difference between true state X_t and the estimated state $\hat{X}_t$. It captures well not only the freshness of information but also the information content of the transmitted packets and the knowledge at the monitor.

The system model and the details on AoII are provided in Chapter 2. Chapter 3 shows the policy that minimizes AoII using Markov Decision Process. The numerical results are provided in Chapter 4.

Chapter 2: System Model

2.1 System Overview

We consider a slotted-time system in which a transmitter sends updates about a process of interest to a remote receiver through an unreliable channel. An illustration of this system is shown in Figure 2.1.

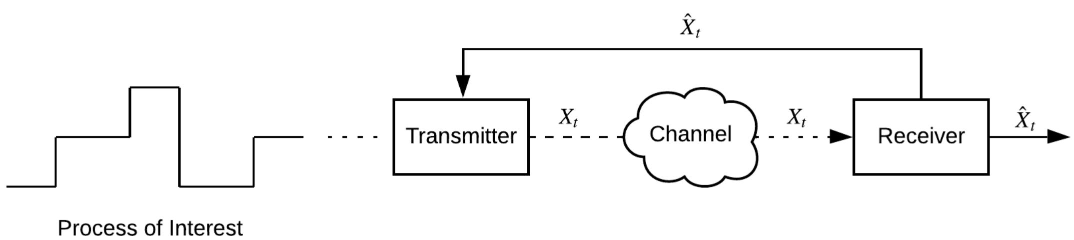

Figure 2.1: Illustration of the system

2.2 Communication Model

The channel is error-free but unreliable. More precisely, the transmission will not necessarily succeed, but if it succeeds, the receiver will receive the exact update the transmitter sent. If the transmission fails, the receiver will receive nothing. We define the channel realization as $r(t)$. $r(t) = 1$ if and only if the transmission is successful and $r(t) = 0$ otherwise. We further define the probability of successful transmission as $Pr(r(t) = 1) = p_s$ and probability of failure transmission as $Pr(r(t) = 0) = 1 - p_s = p_f$. We also assume that $r(t)$ is independent and identically distributed over the time slots. The transmission time for a transmission attempt is deterministic and is equal to the slot duration. For example, if the transmitter schedules a transmission at time t and it succeeds, the receiver will receive the update at time $t + 1$. The failure transmission will also take one slot duration.

As for the transmitter, we assume that the transmitter is capable of generating update X_t by sampling the process at any time on its own will and proceeding to the transmission stage immediately. But the sampling opportunities only occur at the beginning of each time slot. The transmission result will not affect the transmitter's sampling decision. For example, if a transmission happened at time t, regardless of whether the transmission was successful, the transmitter will generate a *new* update at time $t + 1$ if it decides to schedule another transmission.

The receiver will generate a *new* estimate $\hat{X}_t$ every time an update is received. The new estimate will be sent back to the transmitter immediately and received by the transmitter instantaneously. In our model, the receiver uses the received updates as its estimates.

2.3 Age of Incorrect Information

The metric AoII is a penalty function that is defined by true state of the process X_t and the estimated state $\hat{X}_t$. To this end, we assume that X_t and $\hat{X}_t$ are all numerical. The AoII is shown in 2.1.

$$\Delta(t) = \begin{cases} 0 & if \;\; d = 0 \\ \Delta(V(t)) + [t - V(t)] \cdot d & if \;\; d \neq 0 \end{cases} \tag{2.1}$$

where $d = |X_t - \hat{X}_t|$ and $V(t)$ is the last change time of d . When $d = 0$, no penalty is paid since, in this case, the receiver has perfect knowledge of the process. When $d \neq 0$, the penalty will increase depending on the time t and the difference d. Thus, it captures not only the amount of time the estimate has been erroneous but also the distance between the estimate and the true state. A sample path is shown in Figure 2.2.

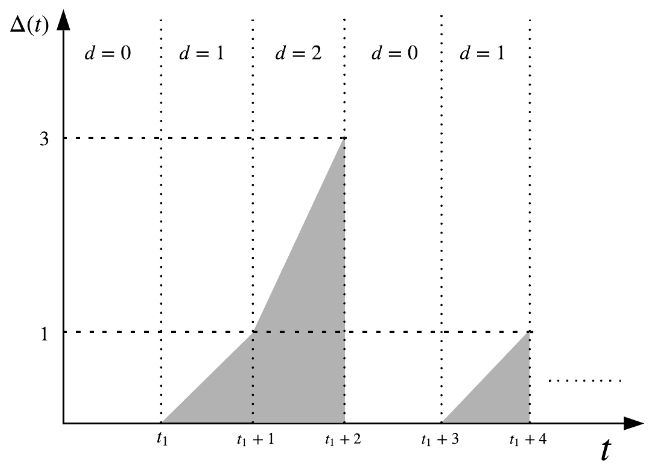

Figure 2.2: A sample path of the penalty

This metric will be the basis of our analysis in the upcoming chapters.

2.4 Source Process

The process of interest is an N-state Markov source where transmissions only happen between adjacent states and with themselves. The corresponding Markov chain is shown in Figure 2.3.

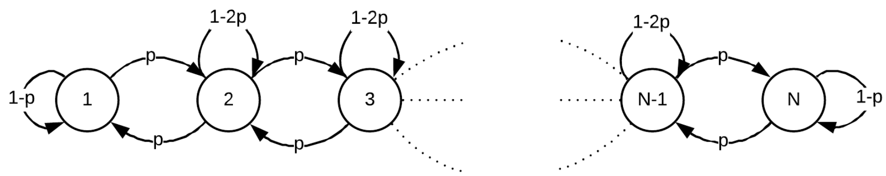

Figure 2.3: N-state Markov source

To simplify our analysis, we ignore the difference caused by the process being at "end-states" (1 and N) and "middle-states" $(2, ..., N-1)$. More precisely, we assume for any $X_t \in \{1, 2, ..., N\}$, $P(X_{t+1} = X_t|X_t) = 1-2p$, $P(X_{t+1} = X_t-1|X_t) = P(X_{t+1} = X_t+1|X_t) = p$. Now, the process dynamics can be fully characterized by the dynamics of the difference d. Thus, we only need to focus on how the difference changes, not on the process's value.

In this case, when no update is received by the receiver (i.e. the estimate does not change), the difference d will not change if the value of the process remains unchanged which happens with probability $1-2p$. When $1 \leq d \leq N-2$, the difference d will either increase or decrease by 1 if the value of the process increases or decreases by 1 which happens with equal probability p. When $d = 0$, the difference d will become 1 if the value of the process changes which happens with probability $2p$. When $d = N-1$, the difference d will become $N-2$ if the value of the process changes which happens with probability $2p$. The corresponding Markov chain is shown in Figure 2.4.

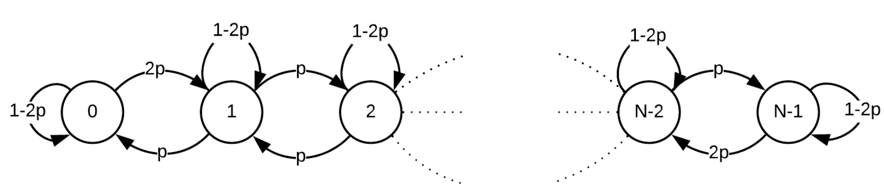

Figure 2.4: Markov chain of the difference d

When an update is received by the receiver (i.e. the estimate changes to the newly received update). The difference d will become 0 if the value of the process remains unchanged which happens with probability $1 - 2p$ and d will become 1 if the value of the process changes which happens with probability $2p$.

In this thesis, we aim to find how the transmitter should act to achieve the minimal expected penalty. To this end, we provide details on the system dynamics in the next section.

2.5 System Dynamics

We define $s(t) \in \mathbb{N}_0$ and $d(t) \in \{0, 1, 2, ..., N - 1\}$ as the penalty and the difference at time t respectively. Then, the system can be characterized by the pair $(d(t), s(t))$. We also define $\psi(t) \in \{0, 1\}$ as the transmitter's decision at time t. $\psi(t) = 1$ means the transmitter decides to schedule a transmission and $\psi(t) = 0$ otherwise. To reduce unnecessary complications, we suppose the system always starts from the state $(0, 0)$. Before characterizing the system dynamics, we first provide the constraints on the pair $(d(t), s(t))$.

- From the penalty function shown in 2.1, $s(t) = 0$ if and only if $d(t) = 0$.
- Since the system always starts from $(0, 0)$ and the penalty either increases or becomes zero at the next time slot, $s(t) \geq \sum_{i=1}^{d(t)} i$. For example, when $d = 2$, the path with minimal penalty is going from $d = 0$ to $d = 2$ via $d = 1$. In this case, $s = \sum_{i=1}^{2} i$. We define $\sum_{i=1}^{d(t)} i = \tau_{d(t)}$.

To summarize, the pair $(d(t), s(t))$ should satisfy the following four constraints:

$$
\begin{aligned}
& d(t) \in \{0, 1, 2, ..., N-1\} \\
& s(t) \in \mathbb{N}_0 \\
& s(t) = 0 \ if \ and \ only \ if \ d(t) = 0 \\
& s(t) \geq \tau_{d(t)} \ where \ \tau_{d(t)} = \sum_{i=1}^{d(t)} i
\end{aligned} \tag{2.2}
$$

Since the system can be fully characterized by the pair $(d(t), s(t))$, we will characterize the values of $(d(t+1), s(t+1))$ using $(d(t), s(t))$ and $\psi(t)$. The dynamics of $d(t)$ can be obtained easily from Section 2.4 and $s(t)$ will change according to 2.1. We distinguish between following cases:

- $(d(t), s(t)) = (0, 0)$: In this case, no matter which action the transmitter takes, the estimate will not change. Thus, $d(t+1) = 0$ with probability $1 - 2p$ and $d(t+1) = 1$ with probability $2p$ as discussed in Section 2.4. Thus, we have:

$$P[(0,0) \mid (0,0)] = (p_s + p_f)(1 - 2p) = 1 - 2p \tag{2.3}$$

$$P[(1,1) \mid (0,0)] = (p_s + p_f)2p = 2p \tag{2.4}$$

- $(d(t), s(t)) = (1, s(t))$ where $s(t) \geq 1$: When $\psi(t) = 0$, from Figure 2.4, we have $d(t+1) = 1$ with probability $1 - 2p$ and $d(t+1) = 0$ or $d(t+1) = 2$ with equal probability p. Thus, we have:

$$P[(0,0) \mid (1, s(t)), \psi = 0] = p \tag{2.5}$$

$$P[(1, s(t)+1) \mid (1, s(t)), \psi = 0] = 1 - 2p \tag{2.6}$$

$$P[(2, s(t)+2) \mid (1, s(t)), \psi = 0] = p \tag{2.7}$$

When $\psi(t) = 1$ and the transmission fails which happens with probability p_f, the dynamics will be the same as those when $\psi(t) = 0$. When the transmission succeeds which happens with probability p_s, $d(t+1) = 0$ with probability $1-2p$ and $d(t+1) = 1$ with probability $2p$ as discussed in Section 2.4. Thus, we have:

$$P[(0,0) \mid (1, s(t)), \psi = 1] = p_s(1-2p) + p_f p \tag{2.8}$$

$$P[(1, s(t)+1) \mid (1, s(t)), \psi = 1] = 2p_s p + p_f(1-2p) \tag{2.9}$$

$$P[(2, s(t)+2) \mid (1, s(t)), \psi = 1] = p_f p \tag{2.10}$$

- $(d(t), s(t))$ where $2 \leq d(t) \leq N-2$ and $s(t) \geq \tau_{d(t)}$: When $\psi(t) = 0$, from Figure 2.4, we have $d(t+1) = d(t)$ with probability $1-2p$ and $d(t+1) = d(t)-1$ or $d(t+1) = d(t)+1$ with equal probability p. Thus, we have:

$$P[(d(t)-1, s(t)+d(t)-1) \mid (d(t), s(t)), \psi = 0] = p \tag{2.11}$$

$$P[(d(t), s(t)+d(t)) \mid (d(t), s(t)), \psi = 0] = 1 - 2p \tag{2.12}$$

$$P[(d(t)+1, s(t)+d(t)+1) \mid (d(t), s(t)), \psi = 0] = p \tag{2.13}$$

When $\psi(t) = 1$ and the transmission fails which happens with probability p_f, the dynamics will be the same as those when $\psi(t) = 0$. When the transmission succeeds which happens with probability p_s, $d(t+1) = 0$ with probability $1-2p$ and $d(t+1) = 1$ with probability $2p$ as discussed in Section 2.4. Thus, we have:

$$P[(0,0) \mid (d(t), s(t)), \psi = 1] = p_s(1-2p) \tag{2.14}$$

$$P[(1, s(t)+1) \mid (d(t), s(t)), \psi = 1] = 2p_s p \tag{2.15}$$

$$P[(d(t)-1, s(t)+d(t)-1) \mid (d(t), s(t)), \psi = 1] = p_f p \tag{2.16}$$

$$P[(d(t), s(t)+d(t)) \mid (d(t), s(t)), \psi = 1] = p_f(1-2p) \tag{2.17}$$

$$P[(d(t)+1, s(t)+d(t)+1) \mid (d(t), s(t)), \psi = 1] = p_f p \tag{2.18}$$

- $(d(t), s(t)) = (N-1, s(t))$ where $s(t) \geq \tau_{N-1}$: When $\psi(t) = 0$, from Figure 2.4, we have $d(t+1) = N-1$ with probability $1-2p$ and $d(t+1) = N-2$ with probability $2p$. Thus, we have:

$$P[(N-1, s(t)+N-1) \mid (N-1, s(t)), \psi = 0] = 1-2p \tag{2.19}$$

$$P[(N-2, s(t)+N-2) \mid (N-1, s(t)), \psi = 0] = 2p \tag{2.20}$$

When $\psi(t) = 1$ and the transmission fails which happens with probability p_f, the dynamics will be the same as those when $\psi(t) = 0$. When the transmission succeeds which happens with probability p_s, $d(t+1) = 0$ with probability $1-2p$ and $d(t+1) = 1$ with probability $2p$ as discussed in Section 2.4. Thus, we have:

$$P[(0,0) \mid (N-1, s(t)), \psi = 1] = p_s(1-2p) \tag{2.21}$$

$$P[(1, s(t)+1) \mid (N-1, s(t)), \psi = 1] = 2p_s p \tag{2.22}$$

$$P[(N-1, s(t)+N-1) \mid (N-1, s(t)), \psi = 1] = p_f(1-2p) \tag{2.23}$$

$$P[(N-2, s(t)+N-2) \mid (N-1, s(t)), \psi = 1] = 2p_f p \tag{2.24}$$

As the system dynamics are fully characterized, we proceed to find how the transmitter should act to achieve the minimal expected penalty.

Chapter 3: Optimal Policy

3.1 Problem Formulation

We aim to find how the transmitter should act to achieve the minimal expected penalty. To this end, we define a series of actions the transmitter takes as $\phi = (\psi(0), \psi(1), ...)$ where $\psi(t)$ is the transmitter's decision at time t. We denote all the feasible series of actions as Φ. Then, the problem of achieving the minimal expected penalty can be formulated as follows:

$$arg \min_{\phi \in \Phi} \lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{T-1} s^{\phi}(t) \mid \phi, (d(0), s(0)) \right] \tag{3.1}$$

where $s^{\phi}(t)$ is the penalty paid at time t when the transmitter acted following the series of actions ϕ. $(d(0), s(0))$ are the initial values of the difference and the penalty respectively. The transmitter should make sequential decisions in a stochastic environment and the penalty depends on decision history.

3.2 MDP Characterization

This minimization problem can be cast into a Markov Decision Process (MDP). Combining the fact that the process never terminates and the object is to achieve the minimal expected penalty, this problem can be further cast into an infinite horizon with average cost MDP that consists of the following components:

- **State**: The state is $K = (d, s)$ where $d \in \{0, 1, ..., N-1\}$ is the difference and s is the penalty.
- **Action**: The feasible action at time t is $\psi(t) \in \{0, 1\}$.
- **Transition Probability**: The transition probabilities between different states of the system are specified in Section 2.5.
- **Cost**: The cost occurred at state (d, s) is simply the penalty s.

Generally, a fixed action sequence won't solve the problem since the system is stochastic. A solution to this problem must specify what the transmitter should do for *any* state the system might reach. The solution of this kind is called a **Policy**. The policy is measured by the expected penalty it achieves and the optimal policy is the one that yields the minimal expected penalty. To find the optimal policy, we first define the loss of state ($V(K)$) which is a single number associated with each state representing the preference of the transmitter. In this case, the transmitter will choose the action that yields the minimal expected loss which is defined as the sum of the loss of all the possible outcomes weighted by the probability of each outcome.

It is well-known that the loss of state can be obtained by solving the Bellman equation [9]. The Bellman equation in infinite horizon with average cost MDP problem is defined as:

$$\theta + V(K) = S_K + \min_{\psi \in \{0,1\}} \left\{ \sum_{K' \in C_K} Pr(K' \mid K, \psi) V(K') \right\} \tag{3.2}$$

where θ is the minimal value of 3.1. $Pr(K' \mid K, \psi)$ is the transition probability from state K to state $K^{'}$ when action ψ is done. C_K is the set of states that are accessible from state K with one transition when ψ is done. S_K is the cost associated with state K.

The Bellman equation can be solved using Value Iteration Algorithm (VIA) [9]. Letting $V_t(\cdot)$ be the loss at iteration t, the Bellman update looks like this:

$$V_{t+1}(K) = S_K + \min_{\psi \in \{0,1\}} \left\{ \sum_{K' \in C_K} Pr(K' \mid K, \psi) V_t(K') \right\} \tag{3.3}$$

VIA is guaranteed to converge to $V(\cdot)$ when $t \to +\infty$ regardless of the initialization (i.e. $\lim_{t \to \infty} V_t(\cdot) = V(\cdot)$). Thus, we will use 3.3 to deduce the optimal policy in the following section.

3.3 Structural Results

Without loss of generality, we assume $V_0(\cdot) = 0$ for all states. We only discuss the pairs that satisfy the constraints in 2.2 in this section. We first provide the following lemma.

Lemma 3.1. *When $p \in [0, \frac{1}{3}]$, the loss of state ($V(d,s)$) is increasing in both d and s (i.e. $V(d, s_1) \geq V(d, s_2)$ $\forall s_1 \geq s_2 \geq 0$ and $V(d_1, s) \geq V(d_2, s)$ $\forall d_1 \geq d_2 \geq 0$)*

Proof. See Appendix A. □

The optimal policy will always choose the action that yields the minimal expected loss. To this end, we define $\Delta V(d,s)$ as the difference between the expected loss at state (d,s) when $\psi = 1$ and $\psi = 0$. More precisely, $\Delta V(d,s) = V^1(d,s) - V^0(d,s)$ where $V^1(d,s)$ and $V^0(d,s)$ are the expected loss at state (d,s) when $\psi = 1$ and $\psi = 0$ respectively.

Theorem 3.2. *When $p \in [0, \frac{1}{3}]$, the optimal policy of the problem in 3.1 is always update policy where the transmitter should schedule transmissions at every time slot or only when $d \neq 0$.*

Proof. See Appendix B. □

Definition 3.3. *Threshold update policy is a policy where the transmitter schedules transmissions only when the current penalty s is greater than or equal to the current threshold. More precisely, when the system is at state (d,s), the transmitter will schedule a transmission only when s is greater than or equal to n_d which is the threshold corresponding to d. In this case, the policy can be fully characterized by the vector $\mathbf{n} = [n_0, n_1, ..., n_{N-1}]$.*

With the above definition provided, we can see that always update policy is a special case of threshold update policy where $\mathbf{n} = \mathbf{0}$. In the next chapter, we will analyze the performance of always update policy by calculating the expected penalty it achieves. As threshold update policy shows potential in more realistic scenario, such as when the transmitter has limited power [8], we will also analyze the performance of threshold update policy in the next chapter.

Chapter 4: Performance

4.1 Always Update Policy

We here evaluate the performance of always update policy by finding the expected penalty it achieves. When this policy is adopted, the transmitter will schedule transmissions at every time slot. In this case, the MDP can be modeled through a Discrete-Time Markov Chain (DTMC). The states refer to the pairs (d, s) where for each $d \in \{1, , , , .N-1\}$, $s \geq \tau_d$ and $s = 0$ if and only if $d = 0$. The transition probabilities can be obtained easily from Section 2.5 and shown in Figure 4.1[1].

In order to find the expected penalty, we start with finding the stationary distribution of the DTMC. Since this DTMC is irreducible, the stationary distribution is well defined. We denote the limiting probability for state (d, s) as $\pi_d(s)$. To make the equations clean and easy to read, we define $\pi_d(s) = 0$ when $0 \leq s < \tau_d$ for $d \neq 0$ and $\pi_0(s) = 0$ when $s > 0$. Then the expected penalty can be calculated as

$$\bar{C} = \sum_{d=0}^{N-1} \left(\sum_{s=0}^{+\infty} s\pi_d(s) \right) = \sum_{d=1}^{N-1} \left(\sum_{s=0}^{+\infty} s\pi_d(s) \right) \tag{4.1}$$

[1]The top arrow of the sub-figure on the upper left corner means that when the system is at state $(N-3, s-N+2)$, it will transfer to $(N-2, s)$ at next time slot with probability $p_f p$. All other arrows can be interpreted in the same way.

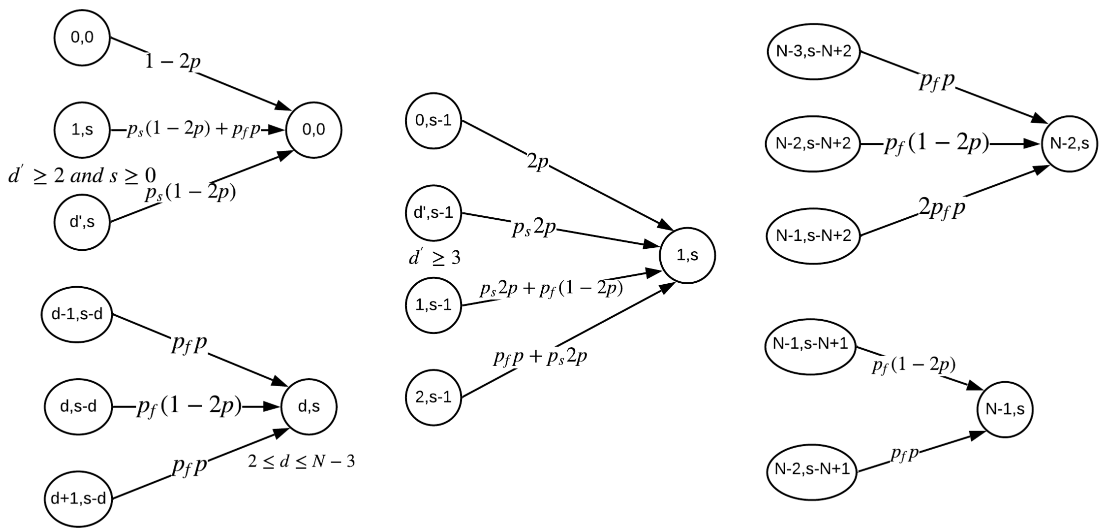

Figure 4.1: Transition probabilities when always update policy is adopted

We define two quantities: $P_d = \sum_{s=0}^{+\infty} \pi_d(s)$ and $C_d = \sum_{s=0}^{+\infty} s\pi_d(s)$. Then the expected penalty can be written as

$$\bar{C} = \sum_{d=1}^{N-1} C_d \tag{4.2}$$

Theorem 4.1. *The expected penalty of always update policy can be obtained by solving the following two systems of linear equations.*

1. system of N+1 linear equations for P_d

$$\sum_{d=0}^{N-1} P_d = 1 \tag{4.3}$$

$$-2pP_0 + p_f p P_1 + [p_s(1-2p)] \sum_{d=1}^{N-1} P_d = 0 \tag{4.4}$$

$$-2pP_0 + (1 + 2p_f p - p_f)P_1 - p_f p P_2 - 2p_s p \sum_{d=1}^{N-1} P_d = 0 \tag{4.5}$$

$$-p_f p P_{N-3} + (1 - p_f + 2p_f p)P_{N-2} - 2p_f p P_{N-1} = 0 \tag{4.6}$$

$$-p_f p P_{N-2} + (1 - p_f + 2p_f p)P_{N-1} = 0 \tag{4.7}$$

For each $d \in \{2, 3, 4, ..., N-3\}$*:*

$$-p_f p P_{d-1} + (1 - p_f + 2p_f p)P_d - p_f p P_{d+1} = 0 \tag{4.8}$$

2. system of N-1 linear equations for C_d

$$(1 + 2p_f p - p_f)C_1 - p_f p C_2 - 2p_s p \sum_{d=1}^{N-1} C_d = P_1 \tag{4.9}$$

$$-p_f p C_{N-3} + (1 - p_f + 2p_f p)C_{N-2} - 2p_f p C_{N-1} = (N-2)P_{N-2} \tag{4.10}$$

$$-p_f p C_{N-2} + (1 - p_f + 2p_f p)C_{N-1} = (N-1)P_{N-1} \tag{4.11}$$

For each $d \in \{2, 3, 4, ..., N-3\}$*:*

$$-p_f p C_{d-1} + (1 - p_f + 2p_f p)C_d - p_f p C_{d+1} = dP_d \tag{4.12}$$

Then, the expected penalty $\bar{C} = \sum_{d=1}^{N-1} C_d$.

Proof. See Appendix C. □

4.2 Threshold Update Policy

When threshold update policy is adopted, the transmitter will schedule transmissions only when the penalty s is greater than or equal to the current threshold. We here consider a case where $n_0 = +\infty$ and $n_{d_1} \geq n_{d_2}\ \forall\ 1 \leq d_1 \leq d_2 \leq N-1$. In this case, the MDP can also be modeled through a Discrete-Time Markov Chain (DTMC). The states refer to the pairs (d, s) where for each $d \in \{1, , , , .N-1\}$, $s \geq \tau_d$ and $s = 0$ if and only if $d = 0$. The transition probabilities can be obtained easily from Section 2.5 and shown in Figure. 4.2[2].

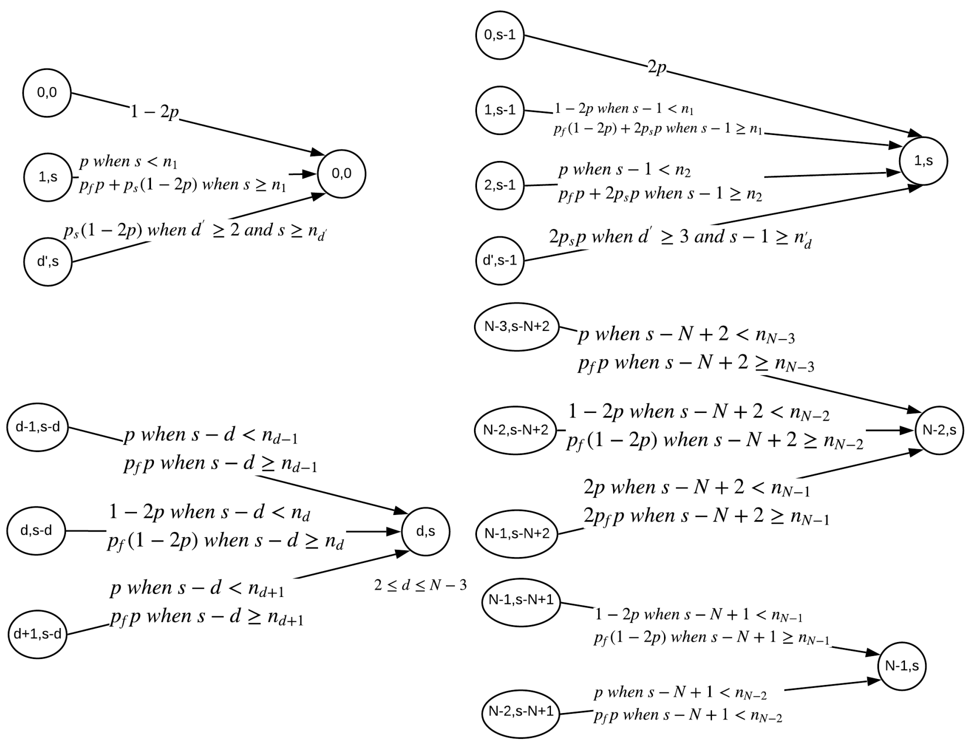

Figure 4.2: Transition probabilities when threshold update policy is adopted

[2]The middle arrow of the sub-figure on the upper right corner means that when the system is at state $(1, s)$, if $s < n_1$, it will transfer to $(0, 0)$ at next time slot with probability p. If $s \geq n_1$, it will transfer to $(0, 0)$ at next time slot with probability $p_f p + p_s(1-2p)$. All other arrows can be interpreted in the same way.

Following the same trajectory as in Section 4.1, the expected penalty in this case can be calculated as

$$\bar{C} = \sum_{d=0}^{N-1} \left(\sum_{s=0}^{+\infty} s\pi_d(s) \right) = \sum_{d=1}^{N-1} \left(\sum_{s=1}^{+\infty} s\pi_d(s) \right) \tag{4.13}$$

We will first give the definitions of some useful quantities. We define $P_d^s = \pi_d(s)$ and $C_d^s = s\pi_d(s)$. We also define $\tau = max\{n_1 + 2, ..., n_{N-2} + N - 1\}$ which is the smallest value such that all the states (d, s) with $d \neq 0$ and $s \geq \tau$ are transferred from the states (d', s') with $s' \geq n_{d'}$. Thus, we can sum all the P_d^s and C_d^s with $s \geq \tau$ separately. More precisely, we define $P_d^{+\infty} = \sum_{s=\tau}^{+\infty} P_d^s$ and $C_d^{+\infty} = \sum_{s=\tau}^{+\infty} C_d^s$. Then the expected penalty can be written as

$$\bar{C} = \sum_{d=1}^{N-1} \left(\sum_{s=1}^{\tau-1} C_d^s + C_d^{+\infty} \right) \tag{4.14}$$

Theorem 4.2. *The expected penalty of threshold update policy can be obtained by solving the following two systems of linear equations.*

1. system of linear equations for P_d^s. There are total of $(N-1)\tau + 2$ linear equations.

First of all, all the probabilities should add up to one.

$$\sum_{d=0}^{N-1} \left(\sum_{s=0}^{\tau-1} P_d^s + P_d^{+\infty} \right) = 1 \tag{4.15}$$

The other equations are shown in Table 4.1.

Linear Equations	When
$0 = -2pP_0^0 + p\sum_{s=1}^{n_1-1} P_1^s + [p_s(1-2p) + p_f p](\sum_{s=n_1}^{\tau-1} P_1^s + P_1^{+\infty})$ $+p_s(1-2p)\sum_{d=2}^{N-1}\{\sum_{s=n_d}^{\tau-1} P_d^s + P_d^{+\infty}\}$	
$0 = 2pP_0^{s-1} - P_1^s + (1-2p)P_1^{s-1} + pP_2^{s-1}$	$s \in \{1, 2, 3, ..., n_{N-1}\}$
$0 = 2pP_0^{s-1} - P_1^s + (1-2p)P_1^{s-1} + pP_2^{s-1} + 2p_s p\sum_{t=d}^{N-1} P_t^{s-1}$	$s \in \{n_d + 1, ..., n_{d-1}\}$ where $3 \leq d \leq N-1$
$0 = 2pP_0^{s-1} - P_1^s + (1-2p)P_1^{s-1} + p_f pP_2^{s-1} + 2p_s p\sum_{t=2}^{N-1} P_t^{s-1}$	$s \in \{n_2 + 1, ..., n_1\}$
$0 = 2pP_0^{s-1} - P_1^s + p_f(1-2p)P_1^{s-1} + p_f pP_2^{s-1} + 2p_s p\sum_{t=1}^{N-1} P_t^{s-1}$	$s \in \{n_1 + 1, ..., \tau - 1\}$
$0 = 2pP_0^{s-1} + p_f(1-2p)P_1^{\tau-1} + (p_f - 2p_f p - 1)P_1^{+\infty} + p_f p(P_2^{\tau-1} + P_2^{+\infty})$ $+2p_s p\sum_{t=1}^{N-1}(P_t^{\tau-1} + P_t^{+\infty})$	
$0 = pP_{N-3}^{s-N+2} - P_{N-2}^s + (1-2p)P_{N-2}^{s-N+2} + 2pP_{N-1}^{s-N+2}$	$s \in \{1, 2, 3, ..., n_{N-1} + N - 3\}$
$0 = pP_{N-3}^{s-N+2} - P_{N-2}^s + (1-2p)P_{N-2}^{s-N+2} + 2p_f pP_{N-1}^{s-N+2}$	$s \in \{n_{N-1} + N - 2, ..., n_{N-2} + N - 3\}$
$0 = pP_{N-3}^{s-N+2} - P_{N-2}^s + p_f(1-2p)P_{N-2}^{s-N+2} + 2p_f pP_{N-1}^{s-N+2}$	$s \in \{n_{N-2} + N - 2, ..., n_{N-3} + N - 3\}$
$0 = p_f pP_{N-3}^{s-N+2} - P_{N-2}^s + p_f(1-2p)P_{N-2}^{s-N+2} + 2p_f pP_{N-1}^{s+N-2}$	$s \in \{n_{N-3} + N - 2, ..., \tau - 1\}$
$0 = p_f p(\sum_{s=\tau-N+2}^{\tau-1} P_{N-3}^s + P_{N-3}^{+\infty}) + p_f(1-2p)\sum_{s=\tau-N+2}^{\tau-1} P_{N-2}^s$ $+(p_f - 2p_f p - 1)P_{N-2}^{+\infty} + 2p_f p(\sum_{s=\tau-N+2}^{\tau-1} P_{N-1}^s + P_{N-1}^{+\infty})$	
$0 = pP_{N-2}^{s-N+1} - P_{N-1}^s + (1-2p)P_{N-1}^{s-N+1}$	$s \in \{1, ..., n_{N-1} + N - 2\}$
$0 = pP_{N-2}^{s-N+1} - P_{N-1}^s + p_f(1-2p)P_{N-1}^{s-N+1}$	$s \in \{n_{N-1} + N - 1, ..., n_{N-2} + N - 2\}$
$0 = p_f pP_{N-2}^{s-N+1} - P_{N-1}^s + p_f(1-2p)P_{N-1}^{s-N+1}$	$s \in \{n_{N-2} + N - 1, ..., \tau - 1\}$
$0 = p_f p(\sum_{s=\tau-N+1}^{\tau-1} P_{N-2}^s + P_{N-2}^{+\infty}) + (p_f - 2p_f p - 1)P_{N-1}^{+\infty}$ $+p_f(1-2p)\sum_{s=\tau-N+1}^{\tau-1} P_{N-1}^s$	
For each $d \in \{2, 3, ..., N-3\}$	
$0 = pP_{d-1}^{s-d} - P_d^s + (1-2p)P_d^{s-d} + pP_{d+1}^{s-d}$	$s \in \{1, 2, 3, ..., n_{d+1} + d - 1\}$
$0 = pP_{d-1}^{s-d} - P_d^s + (1-2p)P_d^{s-d} + p_f pP_{d+1}^{s-d}$	$s \in \{n_{d+1} + d, ..., n_d + d - 1\}$
$0 = pP_{d-1}^{s-d} - P_d^s + p_f(1-2p)P_d^{s-d} + p_f pP_{d+1}^{s-d}$	$s \in \{n_d + d, ..., n_{d-1} + d - 1\}$
$0 = p_f pP_{d-1}^{s-d} - P_d^s + p_f(1-2p)P_d^{s-d} + p_f pP_{d+1}^{s-d}$	$s \in \{n_{d-1} + d, ..., \tau - 1\}$
$0 = p_f p(\sum_{s=\tau-d}^{\tau-1} P_{d-1}^s + P_{d-1}^{+\infty}) + p_f(1-2p)\sum_{s=\tau-d}^{\tau-1} P_d^s + (p_f - 2p_f p - 1)P_d^{+\infty}$ $+p_f p(\sum_{s=\tau-d}^{\tau-1} P_{d+1}^s + P_{d+1}^{+\infty})$	

Table 4.1: System of linear equations for P_d^s

2. system of linear equations for C_d^s. There are total of $(N-1)\tau$ linear equations. The equations are specified in Table 4.2

Linear Equations	When
$-P_1^s = -C_1^s + (1-2p)C_1^{s-1} + pC_2^{s-1}$	$s \in \{1,2,3,...,n_{N-1}\}$
$-P_1^s = -C_1^s + (1-2p)C_1^{s-1} + pC_2^{s-1} + 2p_sp\sum_{t=d}^{N-1} C_t^{s-1}$	$s \in \{n_d+1,...,n_{d-1}\}$ where $3 \le d \le N-1$
$-P_1^s = -C_1^s + (1-2p)C_1^{s-1} + p_fpC_2^{s-1} + 2p_sp\sum_{t=2}^{N-1} C_t^{s-1}$	$s \in \{n_2+1,...,n_1\}$
$-P_1^s = -C_1^s + p_f(1-2p)C_1^{s-1} + p_fpC_2^{s-1} + 2p_sp\sum_{t=1}^{N-1} C_t^{s-1}$	$s \in \{n_1+1,...,\tau-1\}$
$-P_1^{+\infty} = p_f(1-2p)C_1^{\tau-1} + (p_f - 2p_fp - 1)C_1^{+\infty} + p_fp(C_2^{\tau-1} + C_2^{+\infty})$ $+2p_sp\sum_{t=1}^{N-1}(C_t^{\tau-1} + C_t^{+\infty})$	
$-(N-2)P_{N-2}^s = pC_{N-3}^{s-N+2} - C_{N-2}^s + (1-2p)C_{N-2}^{s-N+2} + 2pC_{N-1}^{s-N+2}$	$s \in \{1,2,3,...,n_{N-1}+N-3\}$
$-(N-2)P_{N-2}^s = pC_{N-3}^{s-N+2} - C_{N-2}^s + (1-2p)C_{N-2}^{s-N+2} + 2p_fpC_{N-1}^{s-N+2}$	$s \in \{n_{N-1}+N-2,...,n_{N-2}+N-3\}$
$-(N-2)P_{N-2}^s = pC_{N-3}^{s-N+2} - C_{N-2}^s + p_f(1-2p)C_{N-2}^{s-N+2} + 2p_fpC_{N-1}^{s-N+2}$	$s \in \{n_{N-2}+N-2,...,n_{N-3}+N-3\}$
$-(N-2)P_{N-2}^s = p_fpC_{N-3}^{s-N+2} - C_{N-2}^s + p_f(1-2p)C_{N-2}^{s-N+2} + 2p_fpC_{N-1}^{s-N+2}$	$s \in \{n_{N-3}+N-2,...,\tau-1\}$
$-(N-2)P_{N-2}^{+\infty} = p_fp(\sum_{s=\tau-N+2}^{\tau-1} C_{N-3}^s + C_{N-3}^{+\infty}) + p_f(1-2p)\sum_{s=\tau-N+2}^{\tau-1} C_{N-2}^s$ $+(p_f - 2p_fp - 1)C_{N-2}^{+\infty} + 2p_fp(\sum_{s=\tau-N+2}^{\tau-1} C_{N-1}^s + C_{N-1}^{+\infty})$	
$-(N-1)P_{N-1}^s = pC_{N-2}^{s-N+1} - C_{N-1}^s + (1-2p)C_{N-1}^{s-N+1}$	$s \in \{1,2,3,...,n_{N-1}+N-2\}$
$-(N-1)P_{N-1}^s = pC_{N-2}^{s-N+1} - C_{N-1}^s + p_f(1-2p)C_{N-1}^{s-N+1}$	$s \in \{n_{N-1}+N-1,...,n_{N-2}+N-2\}$
$-(N-1)P_{N-1}^s = p_fpC_{N-2}^{s-N+1} - C_{N-1}^s + p_f(1-2p)C_{N-1}^{s-N+1}$	$s \in \{n_{N-2}+N-1,...,\tau-1\}$
$-(N-1)P_{N-1}^{+\infty} = p_fp(\sum_{s=\tau-N+1}^{\tau-1} C_{N-2}^s + C_{N-2}^{+\infty}) + (p_f - 2p_fp - 1)C_{N-1}^{+\infty}$ $+p_f(1-2p)\sum_{s=\tau-N+1}^{\tau-1} C_{N-1}^s$	
For each $d \in \{2,3,...,N-3\}$	
$-dP_d^s = pC_{d-1}^{s-d} - C_d^s + (1-2p)C_d^{s-d} + pC_{d+1}^{s-d}$	$s \in \{1,2,3,...,n_{d+1}+d-1\}$
$-dP_d^s = pC_{d-1}^{s-d} - C_d^s + (1-2p)C_d^{s-d} + p_fpC_{d+1}^{s-d}$	$s \in \{n_{d+1}+d,...,n_d+d-1\}$
$-dP_d^s = pC_{d-1}^{s-d} - C_d^s + p_f(1-2p)C_d^{s-d} + p_fpC_{d+1}^{s-d}$	$s \in \{n_d+d,...,n_{d-1}+d-1\}$
$-dP_d^s = p_fpC_{d-1}^{s-d} - C_d^s + p_f(1-2p)C_d^{s-d} + p_fpC_{d+1}^{s-d}$	$s \in \{n_{d-1}+d,...,\tau-1\}$
$-dP_d^{+\infty} = p_fp(\sum_{s=\tau-d}^{\tau-1} C_{d-1}^s + C_{d-1}^{+\infty}) + p_f(1-2p)\sum_{s=\tau-d}^{\tau-1} C_d^s + (p_f - 2p_fp - 1)C_d^{+\infty}$ $+p_fp(\sum_{s=\tau-d}^{\tau-1} C_{d+1}^s + C_{d+1}^{+\infty})$	

Table 4.2: System of linear equations for C_d^s

Then the expected penalty $\bar{C} = \sum_{d=1}^{N-1}(\sum_{s=1}^{\tau-1} C_d^s + C_d^{+\infty})$

Proof. The proof is similar to the one in Appendix C . Apart from the increased complexity of calculations, there is no theoretical difference. Thus, the detail is omitted here for the sake of space. □

4.3 Numerical Results

In this section, we provide numerical results concerning the performances of the policies discussed in the previous sections. To this end, we set the initial values of the difference and the penalty to 0. We consider a system where the number of states $N = 7$. All the results are averaged over 100000 time slots.

We first provide numerical results when the transmitter adopts always update policy. We evaluate the effect of process dynamics on the performance of always update policy. We also evaluate the effect of p_s. To this end, for different values of p_s, we vary the probability of changing value (p) and plot the corresponding simulation and theoretical results. The results are shown in Figure 4.3.

We then provide numerical results when the transmitter adopts threshold update policy. The threshold vector $\mathbf{n}$ is chosen randomly as $[+\infty, 15, 13, 10, 7, 5, 3]$. We evaluate the effect of process dynamics as well as the effect of p_s. To this end, for different values of p_s, we vary the probability of changing value (p) and plot the corresponding simulation and theoretical results. The results are shown in Figure 4.4.

We also evaluate the effect of the threshold on the performance of threshold

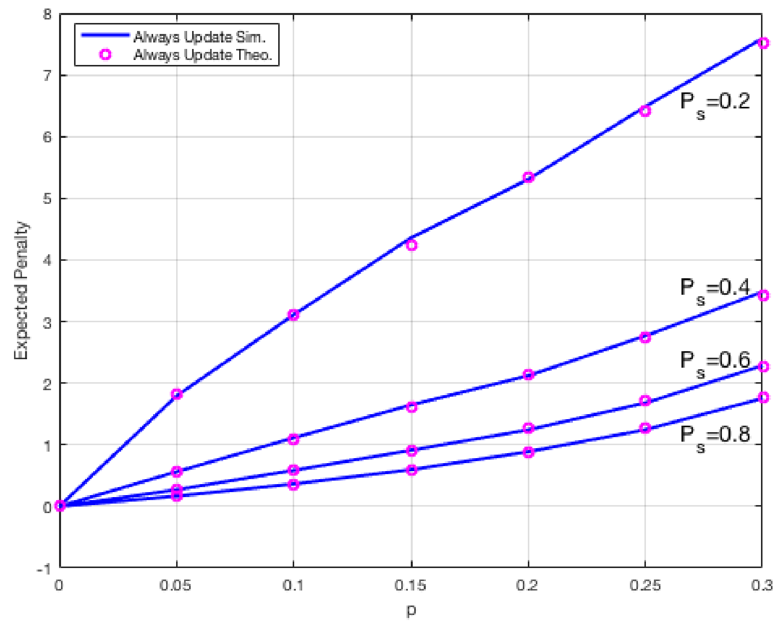

Figure 4.3: Expected penalty as a function of p (always update policy)

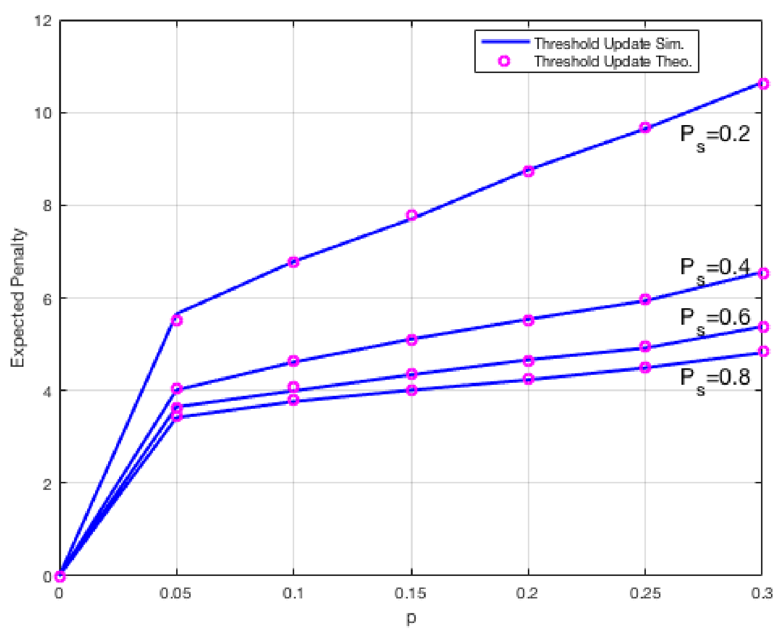

Figure 4.4: Expected penalty as a function of p (threshold update policy)

update policy. To better observe the effect, we consider a simple case where the thresholds are constant which means that $n_d = n$ for $0 \le d \le N-1$. We set the probability of changing value $p = 0.2$. We vary the threshold (n) and plot the corresponding simulation and theoretical results. The results are shown in Figure 4.5.

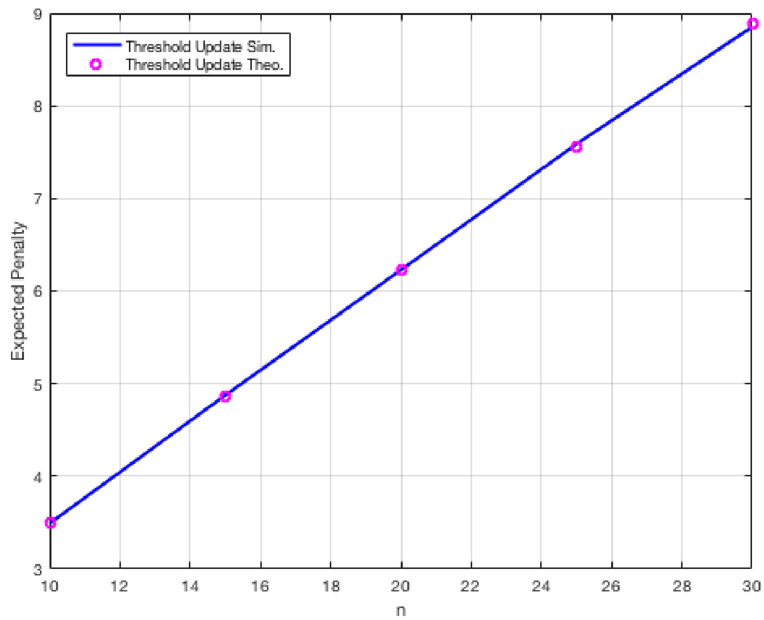

Figure 4.5: Expected penalty as a function of n (threshold update policy)

We can see that, for both policies, the expected penalty increases as p increases. When p increases, the process will more likely change value than remain the same. Then, when the transmission succeeds, the penalty will more likely be one instead of zero. Thus, the penalty will increase as p increases. At the same time, the expected penalty decreases as p_s increases. In our model, failed transmission is equivalent to no transmission which means that the transmission attempts are beneficial only when they are successful. Combining with the fact that the larger p_s

is, the more transmission attempts will succeed, we can conclude that the expected penalty decreases as p_s increases.

For the threshold update policy, the expected penalty will increase as threshold increases. When the threshold is large, the system will allow a large penalty before scheduling a transmission. Thus, the expected penalty will increase.

Chapter 5: Summary and Future Work

5.1 Summary

In this thesis, a new metric - AoII is studied in a system where a transmitter sends updates about an N-state Markov source to a remote receiver through an unreliable channel. Leveraging the MDP tools, it is shown that a simple always update policy minimizes the AoII. A more general transmission policy - threshold update policy is also studied in this thesis. Finally, numerical results are laid out to highlight the effects of system parameters on the performances of both policies.

5.2 Future Work

In the thesis, a simple scenario where there exist no constraints on the capabilities of the transmitter is studied. However, there often exist constraints on the transmitter in real life, such as limited power. Studying AoII under such constraints is of great practical importance. Moreover, the time function in the AoII used is linear. However, in many real-life applications, non-linear increasing time functions, such as quadratic or exponential, will be more reasonable. Thus, AoII with non-linear increasing time function is also worth studying.

Appendix A: Proof of Lemma 3.1

As stated in the lemma, we consider the case where $p \in [0, \frac{1}{3}]$. We know that Value Iteration Algorithm is guaranteed to converge to the solution of Bellman equation when $t \to +\infty$ regardless of the initialization. Thus it is sufficient to prove that $\forall s_1 \geq s_2 \geq 0$

$$V_t(d, s_1) \geq V_t(d, s_2) \tag{A.1}$$

and $\forall d_1 \geq d_2 \geq 0$

$$V_t(d_1, s) \geq V_t(d_2, s) \tag{A.2}$$

A.1 and A.2 hold when $t = 0$ by initialization. We suppose A.1 and A.2 hold up till iteration t. We want to examine whether A.1 and A.2 still hold at iteration $t + 1$. First of all, when $d_1 = d_2$ or $s_1 = s_2$, A.1 and A.2 hold obviously.

We next consider when $s_1 > s_2 > 0$. $\forall s_1 > s_2 > 0$, $V_{t+1}(d, s_1) \geq V_{t+1}(d, s_2)$ holds since the transition probabilities depend only on d and $V_t(d, s)$ is increasing in s.

Now we consider when $d_1 > d_2 > 0$. $\forall d_1 > d_2 > 0$, we examine if $V_{t+1}(d_1, s) \geq V_{t+1}(d_2, s)$. We distinguish between following cases:

- When $d_1 = 2$, $d_2 = 1$:

$$V_{t+1}(2, s) = min\{x, y\} \tag{A.3}$$

where

$$x = s + pV_t(1, s + 1) + (1 - 2p)V_t(2, s + 2) + pV_t(3, s + 3) \tag{A.4}$$

$$\begin{aligned} y \quad = \quad & s + p_s(1 - 2p)V_t(0, 0) + 2p_s pV_t(1, s + 1) \\ & + p_f pV_t(1, s + 1) + p_f(1 - 2p)V_t(2, s + 2) + p_f pV_t(3, s + 3) \end{aligned} \tag{A.5}$$

$$V_{t+1}(1, s) = min\{z, w\} \tag{A.6}$$

where

$$z = s + pV_t(0, 0) + (1 - 2p)V_t(1, s + 1) + pV_t(2, s + 2) \tag{A.7}$$

$$\begin{aligned} w \quad = \quad & s + p_s(1 - 2p)V_t(0, 0) + 2p_s pV_t(1, s + 1) \\ & + p_f pV_t(0, 0) + p_f(1 - 2p)V_t(1, s + 1) + p_f pV_t(2, s + 2) \end{aligned} \tag{A.8}$$

We have

$$\begin{aligned} x - z \quad = \quad & p[V_t(1, s + 1) - V_t(0, 0)] + p[V_t(3, s + 3) - V_t(2, s + 2)] \\ & + (1 - 2p)[V_t(2, s + 2) - V_t(1, s + 1)] \end{aligned} \tag{A.9}$$

Baring in mind that $V_t(d, s)$ is increasing in both d and s, we can easily see that $x - z \geq 0$.

$$y - w = p_f \cdot (x - z) \tag{A.10}$$

Since $p_f \geq 0$, we can easily see that $y - w \geq 0$. Since $V_{t+1}(2, s) = min\{x, y\}$ and $V_{t+1}(1, s) = min\{z, w\}$, we can conclude that $V_{t+1}(2, s) \geq V_{t+1}(1, s)$.

- When $2 \leq d_2 < d_1 \leq N - 2$. In this case, the structures of Bellman update for d_1 and d_2 are the same. Combining the fact that $V_t(d, s)$ is increasing in both d and s, $V_{t+1}(d_1, s) \geq V_{t+1}(d_2, s)$ holds.

- When $d_1 = N - 1$ and $d_2 = N - 2$:

$$V_{t+1}(N - 1, s) = min\{x, y\} \tag{A.11}$$

where

$$x = s + (1 - 2p)V_t(N - 1, s + N - 1) + 2pV_t(N - 2, s + N - 2) \tag{A.12}$$

$$\begin{aligned} y \quad = \quad & s + p_f(1 - 2p)V_t(N - 1, s + N - 1) + 2p_f pV_t(N - 2, s + N - 2) \\ & + p_s(1 - 2p)V_t(0, 0) + 2p_s pV_t(1, s + 1) \end{aligned} \tag{A.13}$$

$$V_{t+1}(N - 2, s) = min\{z, w\} \tag{A.14}$$

where

$$\begin{aligned} z &= s + pV_t(N-3, s+N-3) + pV_t(N-1, s+N-1) \\ &\quad +(1-2p)V_t(N-2, s+N-2) \end{aligned} \tag{A.15}$$

$$\begin{aligned} w &= s + p_s(1-2p)V_t(0,0) + 2p_s pV_t(1, s+1) \\ &\quad +p_f pV_t(N-3, s+N-3) + p_f pV_t(N-1, s+N-1) \\ &\quad +p_f(1-2p)V_t(N-2, s+N-2) \end{aligned} \tag{A.16}$$

We have

$$\begin{aligned} x - z &= (1-3p)V_t(N-1, s+N-1) + (4p-1)V_t(N-2, s+N-2) \\ &\quad -pV_t(N-3, s+N-3) \\ &= (1-3p)[V_t(N-1, s+N-1) - V_t(N-2, s+N-2)] \\ &\quad +p[V_t(N-2, s+N-2) - V_t(N-3, s+N-3)] \end{aligned} \tag{A.17}$$

Baring in mind that $p \in [0, \frac{1}{3}]$ and $V_t(d, s)$ is increasing in both d and s, we can easily see that $x - z \geq 0$.

$$y - w = p_f \cdot (x - z) \tag{A.18}$$

Since $p_f \geq 0$, we can easily see that $y - w \geq 0$. Since $V_{t+1}(N-1, s) = min\{x, y\}$

and $V_{t+1}(N-2,s) = min\{z,w\}$, we can see that $V_{t+1}(N-1,s) \geq V_{t+1}(N-2,s)$.

Till this moment, we have proved that $V_{t+1}(d,s_1) \geq V_{t+1}(d,s_2)$, $\forall s_1 \geq s_2 > 0$ and $V_{t+1}(d_1,s) \geq V_{t+1}(d_2,s)$, $\forall d_1 \geq d_2 > 0$. Next we consider the case where $d_2 = 0$ and $s_2 = 0$. It is sufficient to compare $V_{t+1}(0,0)$ and $V_{t+1}(1,1)$. Thus, we have

$$V_{t+1}(1,1) = min\{x,y\} \tag{A.19}$$

where

$$x = 1 + pV_t(0,0) + (1-2p)V_t(1,2) + pV_t(2,3) \tag{A.20}$$

$$\begin{aligned} y &= 1 + p_s(1-2p)V_t(0,0) + 2p_s pV_t(1,2) \\ &\quad + p_f pV_t(0,0) + p_f(1-2p)V_t(1,2) + p_f pV_t(2,3) \end{aligned} \tag{A.21}$$

$$V_{t+1}(0,0) = min\{z,w\} \tag{A.22}$$

where

$$z = w = (1-2p)V_t(0,0) + 2pV_t(1,1) \tag{A.23}$$

We have

$$\begin{aligned} x - z &= 1 + (3p-1)V_t(0,0) + (1-2p)V_t(1,2) + pV_t(2,3) - 2pV_t(1,1) \\ &= 1 + (1-3p)[V_t(1,2) - V_t(0,0)] + p[V_t(2,3) - V_t(1,1)] \\ &\quad + p[V_t(1,2) - V_t(1,1)] \end{aligned} \tag{A.24}$$

$$y - w = 1 + p_f(1-3p)[V_t(1,2) - V_t(0,0)] + p_f p[V_t(2,3) - V_t(1,2)] + 2p[V_t(1,2) - V_t(1,1)] \quad (A.25)$$

Baring in mind that $p \in [0, \frac{1}{3}]$ and $V_t(d,s)$ is increasing in both d and s, we can easily see that $x - z \geq 0$ and $y - w \geq 0$. Since $V_{t+1}(1,1) = min\{x,y\}$ and $V_{t+1}(0,0) = min\{z,w\}$, we can conclude $V_{t+1}(1,1) \geq V_{t+1}(0,0)$. Now, we have proved that $V(d,s)$ is increasing in both d and s.

Appendix B: Proof of Theorem 3.2

As stated in the theorem, we consider the case where $p \in [0, \frac{1}{3}]$. We define $\Delta V_t(d,s) = V_t^1(d,s) - V_t^0(d,s)$ as the difference between the expected loss when $\psi = 1$ and $\psi = 0$ at iteration t. We will discuss the sign of $\Delta V_t(d,s)$ for each state. To this end, we distinguish between following cases:

- When $d = 0$ and $s = 0$, we have $\Delta V_t(0,0) = 0$. In this case, both decisions are optimal.

- When $d = 1$ and $s \geq 1$, we have $\Delta V_t(1,s) = x - y$ where

$$\begin{aligned} x &= s + p_s(1-2p)V_t(0,0) + 2p_s p V_t(1,s+1) \\ &\quad + p_f p V_t(0,0) + p_f(1-2p)V_t(1,s+1) + p_f p V_t(2,s+2) \end{aligned} \tag{B.1}$$

$$y = s + pV_t(0,0) + (1-2p)V_t(1,s+1) + pV_t(2,s+2) \tag{B.2}$$

After some rearrangements, we have

$$\begin{aligned} \Delta V_t(1,s) &= p_s[(1-3p)V_t(0,0) + (4p-1)V_t(1,s+1) - pV_t(2,s+2)] \\ &= p_s\{(1-3p)[V_t(0,0) - V_t(1,s+1)] \end{aligned}$$

$$+p[V_t(1,s+1)-V_t(2,s+2)]\} \tag{B.3}$$

Since $p_s \geq 0$ and $p \in [0, \frac{1}{3}]$, combining the results in Lemma 3.1, we can conclude that $\Delta V_t(1,s) \leq 0$. Thus, in this case, it is optimal to schedule a transmission.

- When $2 \leq d \leq N-2$ and $s \geq \tau_d$, we have $\Delta V_t(d,s) = x - y$ where

$$\begin{aligned} x &= s + p_s(1-2p)V_t(0,0) + 2p_s pV_t(1,s+1) + p_f pV_t(d-1,s+d-1) \\ &\quad + p_f(1-2p)V_t(d,s+d) + p_f pV_t(d+1,s+d+1) \end{aligned} \tag{B.4}$$

$$y = s + pV_t(d-1,s+d-1) + (1-2p)V_t(d,s+d) + pV_t(d+1,s+d+1) \tag{B.5}$$

After some rearrangements, we have

$$\begin{aligned} \Delta V_t(d,s) &= p_s\{(1-2p)V_t(0,0) + 2pV_t(1,s+1) \\ &\quad -pV_t(d-1,s+d-1) - pV_t(d+1,s+d+1) \\ &\quad -(1-2p)V_t(d,s+d)\} \\ &= p_s\{(1-2p)[V_t(0,0) - V_t(d,s+d)] \\ &\quad +p[V_t(1,s+1) - V_t(d-1,s+d-1)] \\ &\quad +p[V_t(1,s+1) - V_t(d+1,s+d+1)]\} \end{aligned} \tag{B.6}$$

Since $p_s \geq 0$, combining the results in Lemma 3.1, we can conclude that $\Delta V_t(d,s) \leq 0$. Thus, in this case, it is optimal to schedule a transmission.

- When $d = N-1$ and $s \geq \tau_{N-1}$, we have $\Delta V_t(N-1,s) = x - y$ where

$$\begin{aligned} x &= s + p_s(1-2p)V_t(0,0) + 2p_s p V_t(1,s+1) \\ &\quad + p_f(1-2p)V_t(N-1,s+N-1) \\ &\quad + 2p_f p V_t(N-2,s+N-2) \end{aligned} \tag{B.7}$$

$$y = s + (1-2p)V_t(N-1,s+N-1) + 2pV_t(N-2,s+N-2) \tag{B.8}$$

After some rearrangements, we have

$$\begin{aligned} \Delta V_t(N-1,s) &= p_s\{(1-2p)[V_t(0,0) - V_t(N-1,s+N-1)] \\ &\quad + 2p[V_t(1,s+1) - V_t(N-2,s+N-2)]\} \end{aligned} \tag{B.9}$$

Since $p_s \geq 0$, combining the results in Lemma 3.1, we can conclude that $\Delta V_t(N-1,s) \leq 0$. Thus, in this case, it is optimal to schedule a transmission.

Thus,at any iteration t, it is always optimal to schedule a transmission when $d \neq 0$. Combining the fact that, when $d = 0$, either action is optimal, it is optimal for the transmitter to schedule a transmission at every time slot or only when $d \neq 0$. We call such policy as always update policy. Since Value Iteration Algorithm is guaranteed to converge to the solution of Bellman equation, we can conclude that the optimal policy when $p \in [0, \frac{1}{3}]$ is always update policy.

Appendix C: Proof of Theorem 4.1

First of all, all the probabilities should add up to one.

$$\sum_{d=0}^{N-1}\left(\sum_{s=0}^{+\infty}\pi_d(s)\right)=1 \tag{C.1}$$

From Figure 4.1, we have C.2 to C.6

$$\pi_0(0)=(1-2p)\pi_0(0)+p_f p\sum_{s=0}^{+\infty}\pi_1(s)+[p_s(1-2p)]\sum_{d=1}^{N-1}\sum_{s=0}^{+\infty}\pi_d(s) \tag{C.2}$$

For $s \geq 1$

$$\pi_1(s)=2p\pi_0(s-1)+p_f(1-2p)\pi_1(s-1)+p_f p\pi_2(s-1)+2p_s p\sum_{d=1}^{N-1}\pi_d(s-1) \tag{C.3}$$

For $s \geq N-2$

$$\pi_{N-2}(s)=p_f p\pi_{N-3}(s-N+2)+p_f(1-2p)\pi_{N-2}(s-N+2)+2p_f p\pi_{N-1}(s-N+2) \tag{C.4}$$

For $s \geq N-1$

$$\pi_{N-1}(s)=p_f(1-2p)\pi_{N-1}(s-N+1)+p_f p\pi_{N-2}(s-N+1) \tag{C.5}$$

For each $d \in \{2, 3, 4, ..., N-3\}$ and $s \geq d$, we have

$$\pi_d(s) = p_f p \pi_{d-1}(s-d) + p_f(1-2p)\pi_d(s-d) + p_f p \pi_{d+1}(s-d) \tag{C.6}$$

Since $P_d = \sum_{s=0}^{+\infty} \pi_d(s)$ and $\pi_0(s) = 0$ when $s > 0$, the following equations always hold

$$\pi_0(0) = \sum_{s=0}^{+\infty} \pi_0(s) = P_0 \tag{C.7}$$

Since $\pi_d(s) = 0$ when $0 \leq s < \tau_d$ for $d \neq 0$, when $d \in \{1, 2, ..., N-1\}$, the following equations always hold

$$\sum_{s=\tau_d}^{+\infty} \pi_d(s) = \sum_{s=d}^{+\infty} \pi_d(s) = \sum_{s=0}^{+\infty} \pi_d(s) = P_d \tag{C.8}$$

According to the definition of P_d, C.1 can be written as

$$\sum_{d=0}^{N-1} P_d = 1 \tag{C.9}$$

Using C.7 and C.8, C.2 can be written as

$$P_0 = (1-2p)P_0 + p_f p P_1 + [p_s(1-2p)] \sum_{d=1}^{N-1} P_d \tag{C.10}$$

After some rearrangements, we have 4.4

From C.3, we have

$$\begin{aligned}\sum_{s=1}^{+\infty}\pi_1(s) &= \sum_{s=1}^{+\infty}\{2p\pi_0(s-1)+p_f(1-2p)\pi_1(s-1)\\ &\quad +p_fp\pi_2(s-1)+2p_sp\sum_{d=1}^{N-1}\pi_d(s-1)\}\\ &= \sum_{s=0}^{+\infty}\{2p\pi_0(s)+p_f(1-2p)\pi_1(s)+p_fp\pi_2(s)+2p_sp\sum_{d=1}^{N-1}\pi_d(s)\}\end{aligned} \tag{C.11}$$

Using C.7 and C.8, we have

$$P_1 = 2pP_0 + p_f(1-2p)P_1 + p_fpP_2 + 2p_sp\sum_{d=1}^{N-1}P_d \tag{C.12}$$

After some rearrangements, we have 4.5

From C.4, we have

$$\begin{aligned}\sum_{s=N-2}^{+\infty}\pi_{N-2}(s) &= \sum_{s=N-2}^{+\infty}\{p_fp\pi_{N-3}(s-N+2)\\ &\quad +p_f(1-2p)\pi_{N-2}(s-N+2)+2p_fp\pi_{N-1}(s-N+2)\}\\ &= \sum_{s=0}^{+\infty}\{p_fp\pi_{N-3}(s)+p_f(1-2p)\pi_{N-2}(s)+2p_fp\pi_{N-1}(s)\}\end{aligned} \tag{C.13}$$

Using C.8, we have

$$P_{N-2} = p_fpP_{N-3} + p_f(1-2p)P_{N-2} + 2p_fpP_{N-1} \tag{C.14}$$

After some rearrangements, we have 4.6

From C.5, we have

$$\begin{aligned}\sum_{s=N-1}^{+\infty} \pi_{N-1}(s) &= \sum_{s=N-1}^{+\infty} \{p_f(1-2p)\pi_{N-1}(s-N+1) + p_f p \pi_{N-2}(s-N+1)\} \\ &= \sum_{s=0}^{+\infty} \{p_f(1-2p)\pi_{N-1}(s) + p_f p \pi_{N-2}(s)\} \end{aligned} \tag{C.15}$$

Using C.8, we have

$$P_{N-1} = p_f(1-2p)P_{N-1} + p_f p P_{N-2} \tag{C.16}$$

After some rearrangements, we have 4.7.

When $d \in \{2, 3, 4, ..., N-3\}$, from C.6, we have

$$\begin{aligned}\sum_{s=d}^{+\infty} \pi_d(s) &= \sum_{s=d}^{+\infty} \{p_f p \pi_{d-1}(s-d) + p_f(1-2p)\pi_d(s-d) + p_f p \pi_{d+1}(s-d)\} \\ &= \sum_{s=0}^{+\infty} \{p_f p \pi_{d-1}(s) + p_f(1-2p)\pi_d(s) + p_f p \pi_{d+1}(s)\} \end{aligned} \tag{C.17}$$

Using C.8, we have

$$P_d = p_f p P_{d-1} + p_f(1-2p)P_d + p_f p P_{d+1} \tag{C.18}$$

After some arrangements, we have 4.8

Thus, to solve P_d, we have the system of linear equations shown in Theorem 4.1. This is a system of $N+1$ linear equations.

C_d can be solved in a very similar way. For $d \in \{1, 2, ..., N-1\}$, the following equations always hold

$$\sum_{s=\tau_d}^{+\infty} s\pi_d(s) = \sum_{s=d}^{+\infty} s\pi_d(s) = \sum_{s=0}^{+\infty} s\pi_d(s) = C_d \tag{C.19}$$

Multiplying $s-1$ to the both sides of C.3 and summing over s from 1 to $+\infty$, we have

$$\begin{aligned}\sum_{s=1}^{+\infty}(s-1)\pi_1(s) &= \sum_{s=1}^{+\infty}(s-1)\{2p\pi_0(s-1)+p_f(1-2p)\pi_1(s-1)+p_f p\pi_2(s-1) \\ &\quad +2p_s p\sum_{d=1}^{N-1}\pi_d(s-1)\}\end{aligned} \tag{C.20}$$

After some rearrangements, we have

$$\sum_{s=1}^{+\infty} s\pi_1(s) - \sum_{s=1}^{+\infty}\pi_1(s) = \sum_{s=0}^{+\infty} s\{2p\pi_0(s)+p_f(1-2p)\pi_1(s)+p_f p\pi_2(s)+2p_s p\sum_{d=1}^{N-1}\pi_d(s)\} \tag{C.21}$$

Using C.8 and C.19, we have

$$C_1 - P_1 = p_f(1-2p)C_1 + p_f pC_2 + 2p_s p\sum_{d=1}^{N-1} C_d \tag{C.22}$$

Then, we have 4.9

Multiplying $s-N+2$ to the both sides of C.4 and summing over s from $N-2$ to $+\infty$, we have

$$\sum_{s=N-2}^{+\infty}(s-N+2)\pi_{N-2}(s) = \sum_{s=N-2}^{+\infty}(s-N+2)\{p_f p\pi_{N-3}(s-N+2)$$

$$+p_f(1-2p)\pi_{N-2}(s-N+2)$$

$$+2p_fp\pi_{N-1}(s-N+2)\} \tag{C.23}$$

After some rearrangements, we have

$$\sum_{s=N-2}^{+\infty} s\pi_{N-2}(s) - \sum_{s=N-2}^{+\infty}(N-2)\pi_{N-2}(s) = \sum_{s=0}^{+\infty} s\{p_fp\pi_{N-3}(s) + p_f(1-2p)\pi_{N-2}(s) + 2p_fp\pi_{N-1}(s)\} \tag{C.24}$$

Using C.8 and C.19, we have

$$C_{N-2} - (N-2)P_{N-2} = p_fpC_{N-3} + p_f(1-2p)C_{N-2} + 2p_fpC_{N-1} \tag{C.25}$$

Then, we have 4.10

Multiplying $s-N+1$ to the both sides of C.5 and summing over s from $N-1$ to $+\infty$, we have

$$\sum_{s=N-1}^{+\infty}(s-N+1)\pi_{N-1}(s) = \sum_{s=N-1}^{+\infty}(s-N+1)\{p_f(1-2p)\pi_{N-1}(s-N+1) + p_fp\pi_{N-2}(s-N+1)\} \tag{C.26}$$

After some rearrangements, we have

$$\sum_{s=N-1}^{+\infty} s\pi_{N-1}(s) - \sum_{s=N-1}^{+\infty}(N-1)\pi_{N-1}(s) = \sum_{s=0}^{+\infty} s\{p_f(1-2p)\pi_{N-1}(s) + p_fp\pi_{N-2}(s)\} \tag{C.27}$$

Using C.8 and C.19, we have

$$C_{N-1} - (N-1)P_{N-1} = p_f(1-2p)C_{N-1} + p_f p C_{N-2} \tag{C.28}$$

Then, we have 4.11

For $2 \leq d \leq N-3$, multiplying $s-d$ to the both sides of C.6 and summing over s from d to $+\infty$, we have

$$\sum_{s=d}^{+\infty}(s-d)\pi_d(s) = \sum_{s=d}^{+\infty}(s-d)\{p_f p\pi_{d-1}(s-d) + p_f(1-2p)\pi_d(s-d) + p_f p\pi_{d+1}(s-d)\} \tag{C.29}$$

After some rearrangements, we have

$$\sum_{s=d}^{+\infty} s\pi_d(s) - \sum_{s=d}^{+\infty} d\pi_d(s) = \sum_{s=0}^{+\infty} s\{p_f p\pi_{d-1}(s) + p_f(1-2p)\pi_d(s) + p_f p\pi_{d+1}(s)\} \tag{C.30}$$

Using C.8 and C.19, we have

$$C_d - dP_d = p_f pC_{d-1} + p_f(1-2p)C_d + p_f pC_{d+1} \tag{C.31}$$

Then, we have 4.12

Thus, to solve C_d, we have the system of linear equations shown in Theorem 4.1. This is a system of $N-1$ linear equations.

Bibliography

[1] S. Kaul, R. Yates, and M. Gruteser. Real-time status: How often should one update? In *2012 Proceedings IEEE INFOCOM*, pages 2731–2735, 2012.

[2] S. K. Kaul, R. D. Yates, and M. Gruteser. Status updates through queues. In *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–6, 2012.

[3] C. Kam, S. Kompella, and A. Ephremides. Age of information under random updates. In *2013 IEEE International Symposium on Information Theory*, pages 66–70, 2013.

[4] M. Costa, M. Codreanu, and A. Ephremides. Age of information with packet management. In *2014 IEEE International Symposium on Information Theory*, pages 1583–1587, 2014.

[5] Y. Hsu, E. Modiano, and L. Duan. Age of information: Design and analysis of optimal scheduling algorithms. In *2017 IEEE International Symposium on Information Theory (ISIT)*, pages 561–565, 2017.

[6] Igor Kadota, Abhishek Sinha, Elif Uysal-Biyikoglu, Rahul Singh, and Eytan Modiano. Scheduling policies for minimizing age of information in broadcast wireless networks, 2018.

[7] Ali Maatouk, Saad Kriouile, Mohamad Assaad, and Anthony Ephremides. On the optimality of the whittle's index policy for minimizing the age of information, 2020.

[8] Ali Maatouk, Saad Kriouile, Mohamad Assaad, and Anthony Ephremides. The age of incorrect information: A new performance metric for status updates, 2019.

[9] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall Press, USA, 3rd edition, 2009.