# Age of Information Minimization in Goal-Oriented Communication with Processing and Cost of Actuation Error Constraints

**Rishabh Pomaje[1], Jayanth S[2], Rajshekhar V. Bhat[3], Nikolaos Pappas[4]**

[1, 3]Department of Electrical, Electronics, and Communication Engineering,
Indian Institute of Technology Dharwad, India.
Email: 210020036@iitdh.ac.in, rajshekhar.bhat@iitdh.ac.in@iitdh.ac.in

[2] Nokia Solutions and Networks, Bengaluru, India.
Email: jayanth.s@nokia.com

[4] Department of Computer and Information Science,
Linköping University, Sweden.
Email: nikolaos.pappas@liu.se

August 12, 2025

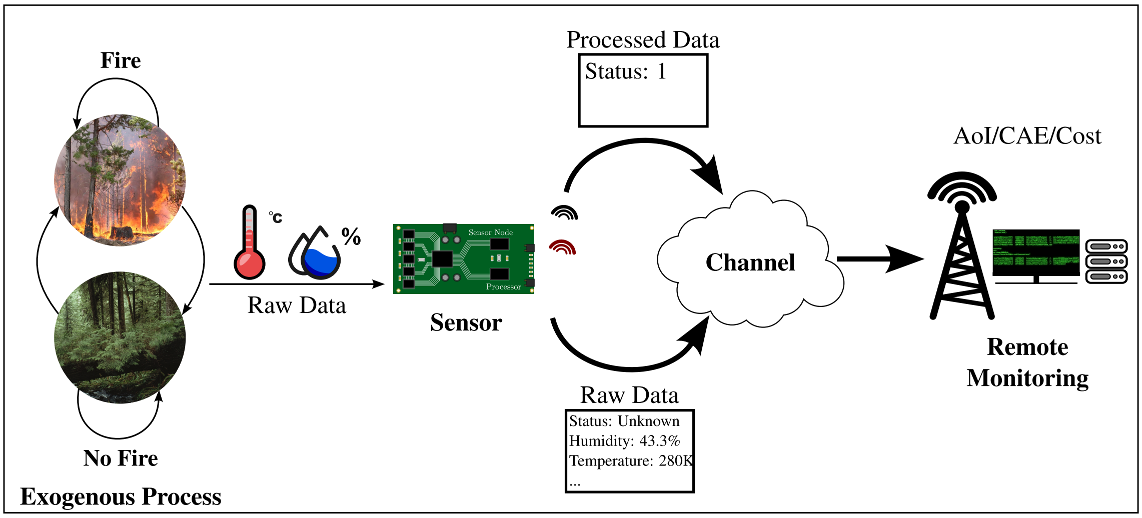



Figure 1: A graphical abstract of the problem addressed in this work.

# Contents


**Chapter 1** **The Problem and a Solution** — Page 4

1.1 Motivation — 4

1.2 System Model — 5

Source Model — 5 • Possible Actions at the Source — 6 • Channel Model — 7 • Cost of Actuation Error - A Context-Aware Distortion Measure — 7 • Age of Information Model — 9

1.3 Optimization Analysis — 10

The General Problem — 10 • Stationary Randomized Policies (SRPs) — 10

**Chapter 2** **Numerical Analysis** — Page 14

Feasibility Analysis — 14 • Impact of Action Costs — 15 • Tradeoffs Between AoI and CAE — 16

**Chapter 3** **Simulation Methodology** — Page 18

3.1 Information Distortion — 18

Simulation Methodology — 19 • Closeness Metrics — 19

3.2 Results and Analysis — 20

**Chapter 4** **A Swarm of Sensors** — Page 23

4.1 System Model — 24

Sources — 24 • Channel — 24 • Possible Decisions by the Controller — 25 • Cost of Actuation Error — 26 • Age of Information — 26

4.2 Optimization Analysis — 26

Problem Formulation — 26 • Real-time expected CAE expression — 27

**Chapter 5** **Appendix** — Page 28

5.1 Distributions of the State and Estimate — 28

5.2 Lower Bound on AoI-metric — 29

5.3 Long-term time-average of expected AoI under SRP — 34

Transition Probabilities — 35

# Notation

The following notation will be followed throughout the text unless an exception is made. The exception will be duly mentioned wherever made.

## General variables

| | |
|---|---|
| $a$ | A real scalar |
| $a(n)$ | The state of a random process a($\cdot$) at time $n$ |
| $A$ | An event |
| $\bar{A}$ | Long-term time average of expectation of process $a(n)$ |
| $\boldsymbol{a}$ | A vector |
| $\boldsymbol{A}$ | A matrix |
| $\mathcal{A}$ | A set |
| $n$ | General time slot index |
| $p$ | Probability of some event that will be clear from the context |
| $P(\cdot)$ | Probability of an event |
| $\mathbb{E}[\cdot]$ or $\mathbb{E}\{\cdot\}$ | Expectation operator |
| $\{x(1), x(2), x(3)\}$ | A set containing $x(1), x(2),$ and $x(3)$ |
| $[x(1), x(2), x(3)]$ | Inline representation of a column vector containing the elements $x(1), x(2),$ and $x(3)$ |

## Special meaning variables

| | |
|---|---|
| $p_{i,j}^{M}$ | Probability of transitioning from state $i$ to state $j$ in a Markov chain $M$ |
| $\boldsymbol{P}_M$ | State-transition matrix of Markov chain $M$ |
| $\boldsymbol{\pi}_M$ | Stationary-Distribution of Markov Chain $M$ |
| $\pi_i^M$ | Probability of being in state $i$ of Markov Chain $M$ under stationarity |

## Reserved variables

| | |
|---|---|
| $A(n), \bar{A}$ | Age of information and its long-term time average of expectation |
| $\Delta(n), \bar{\Delta}$ | Distortion and its long-term time average of expectation |
| $C_{i,j}$ | Cost of Actuation error (Distortion) when true state is $i$ and estimate is $j$ |
| $C(n), \bar{C}$ | Cost of Actions and its long-term time average of expectation |
| $\mathcal{P}$ | Class of general non-anticipative policies |
| $\mathcal{P}_R$ | Class of Stationary Randomized Policies |
| $\mathsf{P}$ | A policy of general class |
| $\mathsf{R}$ | A policy of SRP class |

# Chapter 1

# The Problem and a Solution

## 1.1 Motivation

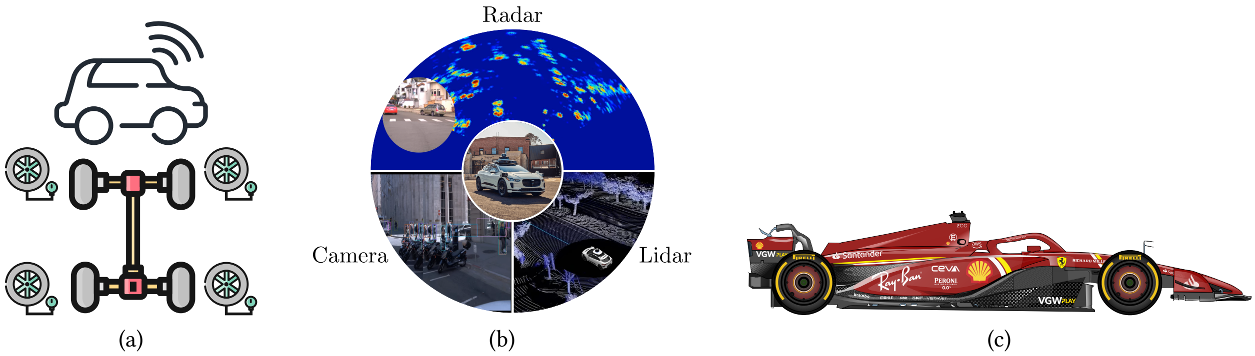



Figure 1.1: (a) Wireless tire pressure monitoring system (less than 10 sensors), (b) Autonomous vehicles/ Intelligent transport (less than 100 sensors). Systems such as shown above consist of numerous sub-systems such as radar, lidar, camera, and other sensors that coordinate and take real-time decisions. Image credits: https://waymo.com/, and (c) Large heterogenous sensor network on an F1 car (more than 300 sensors). To put into perspective, a modern F1 car produces 1.1 million data points every second.

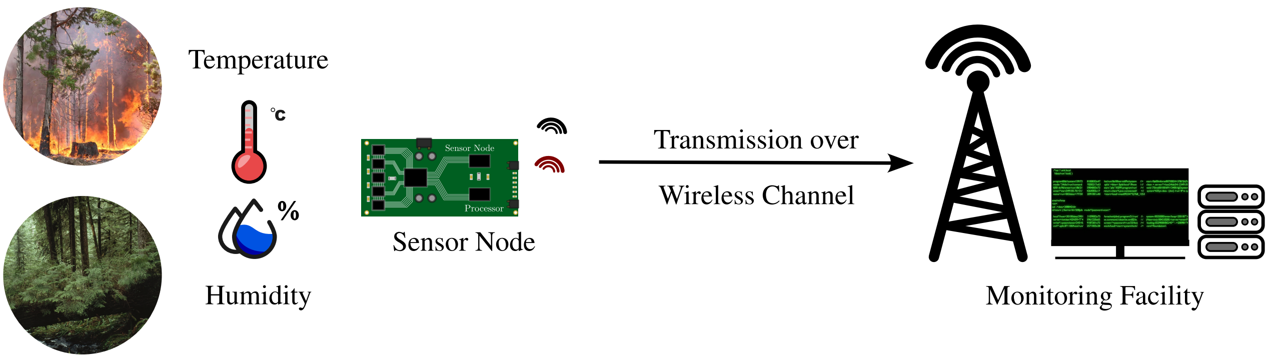



Figure 1.2: Several applications involve monitoring two-state processes. In this example, an off-site agency/ facility wishes to monitor a forest area for wildfires. In this case, Fire/ No Fire can be considered as the two states. To update the facility, the transmitter can send either the raw sample (photo/geo-parameters) or just the state itself after processing the sample locally.

Delivery of fresh information is a key characteristic of the next generation networks and communication systems. The freshness of information is quantified using a metric, "Age of Information". See [Yates et al., 2021, Kaul et al., 2012] for original discussion. This metric measures the time elapsed between the creation of the update and its delivery. Age is an essential characteristic of data especially in use cases such as remote monitoring of some environment. Cyber-physical systems such as forest fire monitoring, concentration-monitoring in a chemical-based factory unit, and real-time embedded systems such as autonomous driving (see figure 1.1) are some instances. Additionally, these systems involve taking decisions based on the state of some source. That it, these systems are said to be **semantics-** or **context-aware**. Based on the estimate of the state of an external process, certain actions are required on the part of the system. As a simple example, in case there is a fire somewhere, the extinguishing procedure needs to be enabled as soon as possible.

In this work, we consider a system where a facility remotely monitors an external, stochastic phenomenon and is interested in being informed about its state. For example[1], consider a forest-fire monitoring system (see figure 1.2). Here, the system must sample the environment and process the sample to estimate its state. The sample can be either parameters such as the humidity, temperature or even an image of the surveillance area. The processing[2] can be done either at the source itself or at the destination. This decision of where to sample is made non-trivial by the asymmetry in costs of transmission and processing. Processing at the source will incur higher costs as compared to processing at the facility, which may be equipped with servers. However, the amount of data to be transmitted would be smaller if the sample was processed prior to transmission and thereby less vulnerable to channel impairments. Thus, a processed sample would offer higher *deliverability*.

Considering all these factors, we wish to design a policy that will deliver fresh information to the facility under constraints on action-dependent and distortion-dependent costs. We begin by describing each part of the system model.

## 1.2 System Model

Let the time be slotted, indexed as $n = 1, 2, \ldots, N, \ldots$. In every slot, the destination (monitoring facility) must decide on an action and convey it to the source (sensor node). The source acts on the decision as instructed and if the destination demands a sample, the corresponding packet transmitted by the source could contain the raw or processed data, also specified by the destination. We assume that the complete process of conveying a decision, generating a sample, transmission, and estimation of information at the destination, for any type of sample, takes place within a single slot. Refer to figure 1.3 for the complete system model.

**Note:-**

The destination by virtue of the setup can only estimate the state of the environment based on the packets it receives from the source. As a result, it may be safely concluded that any policy, including the optimal one, shall be *state-independent*. In a slot, if a transmission does not occur/is unsuccessful, the current estimate carries over as the next estimate.

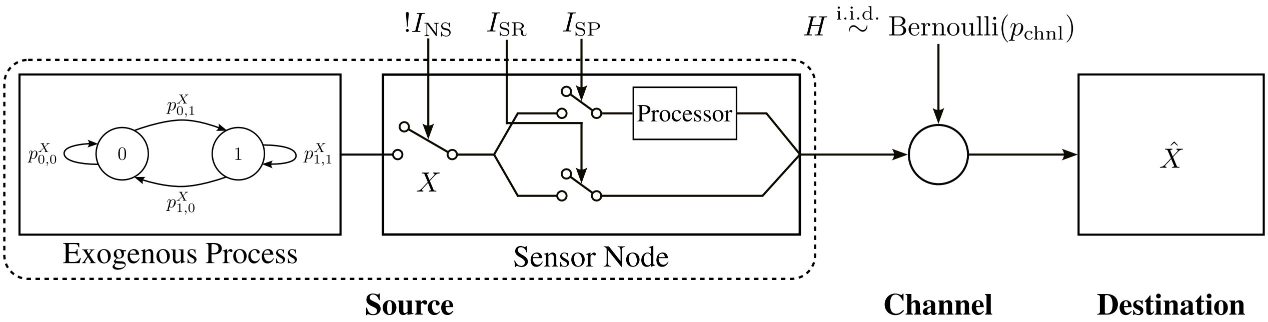



Figure 1.3: The source, directed by the destination, samples the stochastic process on demand and transmits data, processed or unprocessed, over a wireless channel.

### 1.2.1 Source Model

[1] We shall refer to this example throughout this work to justify definitions or develop intuition.

[2] As an example, if the system generates images of the area, a computer vision algorithm can used to provide a distribution over the possible states fire/no-fire.

The exogenous two-state process we are interested in monitoring will be modeled as a discrete-time, time-homogeneous Markov chain (DTMC), whose states are $X(1), X(2), \ldots$. Let $\mathcal{X} = \{0, 1\}$ be its state space, and let the probability of $X(n)$ transitioning from State $i$ to State $j$ be $p^X_{i,j}$ for any $n$. We denote the stationary distribution of this two-state DTMC as $\boldsymbol{\pi}_X = [\pi^X_0, \pi^X_1]$ where,

$$\pi^X_0 = \frac{p^X_{1,0}}{p^X_{0,1} + p^X_{1,0}} \text{ and } \pi^X_1 = \frac{p^X_{0,1}}{p^X_{0,1} + p^X_{1,0}}.$$

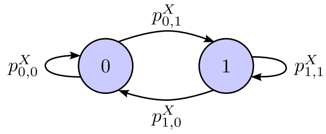



Figure 1.4: A simple 2-state DTMC.

### 1.2.2 Possible Actions at the Source

At the start of each time slot, the source can take one of three possible actions:

1. remain idle (*no sample*, denoted as NS),
2. sample and transmit a *raw* observation (SR),
3. sample and *process* the data before transmission (SP).

We shall refer to the action by their abbreviation in the rest of the text. Consider the following definitions:

**Definition 1.2.1: Action Set, Indicator Variable Function**

Let action set $\mathcal{A}$ be the set of all the possible actions available at the source. Thus, $\mathcal{A} = \{\text{NS}, \text{SR}, \text{SP}\}$. If $I_A(n)$ is used to indicate whether an $A \in \mathcal{A}$ is chosen during slot $n$ then,

$$I_A(n) = \begin{cases} 1 & \text{if the source takes action } A \text{ in slot } n, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

In each time slot $n$, the source will perform one action from $\mathcal{A}$. In words, if $I_{\text{SR}}(n) = 1$ or $I_{\text{SP}}(n) = 1$, the source will send the information packet, raw sample or state, respectively.

Each of the actions from the set $\mathcal{A}$ has a cost[3][4] associated with it. Let $C(n)$ denote the instantaneous transmission cost function. Then, $C(n) = \boldsymbol{c}^\top \boldsymbol{I}(n)$, where $\boldsymbol{c} = [c_{\text{NS}}, c_{\text{SR}}, c_{\text{SP}}]$ is the vector of costs associated with each of the possible actions, and $\boldsymbol{I}(n)$ is the action vector for slot $n$ defined as $\boldsymbol{I}(n) = [I_{\text{NS}}(n), I_{\text{SR}}(n), I_{\text{SP}}(n)]$. We consider the following constraint on the long-term time-averaged expected transmission cost, denoted $\bar{C}$, which is given by

$$\bar{C} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[C(n)] \leqslant c_0,$$

where $c_0$ is the maximum allowed limit on $\bar{C}$ and the expectation is with respect to the action vector.

**Note:-**

One can interpret $c_{\text{NS}}$ as the cost associated with keeping the system running, even when no sample is generated. The costs for sampling and transmitting the raw sample, or sampling, processing, and then transmitting, are captured by $c_{\text{SR}}$ and $c_{\text{SP}}$, respectively. While one could have set $c_{\text{NS}} = 0$ and subtracted this baseline cost from $c_0$, the problem formulation would remain unchanged. However, we explicitly retain $c_{\text{NS}}$ to align with the probabilities $[p^{\mathsf{R}}_{\text{NS}}, p^{\mathsf{R}}_{\text{SR}}, p^{\mathsf{R}}_{\text{SP}}]$, which correspond to not sampling, sampling and transmitting, and sampling, processing, and transmitting, respectively. These probabilities define the stationary randomized policy we adopt as a solution to the problem that will be soon formulated.

[3] The cost associated with transmission actions is distinct from CAE (to be introduced later). This cost, referred to as the transmission cost, arises from the operation of the transmitter and depends on the specific action taken—whether a sample is transmitted or not, and if transmitted, whether it is in raw or processed form. Thus, it may or may also include the sampling and/or processing cost.

[4] We ignore the cost incurred at the destination as we consider them to be negligible compared to the resources available.

### 1.2.3 Channel Model

We consider a single-hop, wireless channel that may be in a *good* or *bad* state. In any given time slot, the channel is considered to be in a good state with probability $p_{\text{chnl}} \in [0,1]$. *Ceteris paribus*, a good channel state is typically associated with a higher probability of successful packet delivery. The random variable $H(n)$ will be used to represent the state of the channel in slot $n$. $H(n) = 1$ would mean that a good channel was realized during slot $n$. Thus,

$$H(n) \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p_{\text{chnl}}), \forall\, n. \tag{1.2}$$

The probability that the packet is successfully delivered, denoted by $p_h^t \in [0,1]$, depends on the channel condition $h$ and type of packet $t$ for all $h \in \{0\ (\text{bad}), 1\ (\text{good})\}$ and $t \in \{r\ (\text{raw}), p\ (\text{processed})\}$.

A decision rule or a policy would specify the action vector, $\boldsymbol{I}(n)$, for all $n$. Since the source has to perform at least and at most one action from $\mathcal{A}$, any feasible policy must satisfy the following *validity constraints*:

$$I_A(n) \cdot (I_A(n) - 1) = 0 \quad \forall\, A \in \mathcal{A}, n \in \mathbb{N}, \tag{1.3a}$$

$$\sum_{\mathcal{A}} I_A(n) = 1 \quad \forall\, n \in \mathbb{N}. \tag{1.3b}$$

Let $d(n)$ be the random variable indicating whether a packet of information is successfully *delivered* to the destination during slot $n$. It takes the value 1 if the packet is successfully delivered and 0 otherwise. Hence,

$$d(n) = \begin{cases} 1, & \text{with probability } \mu, \text{if } \boldsymbol{I}(n) = [0,1,0], \\ 1, & \text{with probability } \nu, \text{if } \boldsymbol{I}(n) = [0,0,1], \\ 0, & \text{with probability} 1, \text{otherwise,} \end{cases}$$

where $\nu = p_0^p \cdot p_{\text{chnl}} + p_0^p \cdot (1 - p_{\text{chnl}})$ and $\mu = p_0^r \cdot p_{\text{chnl}} + p_0^r \cdot (1 - p_{\text{chnl}})$. Thus, the conditional expectation of $d(n)$ is given by

$$\mathbb{E}[d(n) \mid \boldsymbol{I}(n)] = \mu \cdot I_{\text{SR}}(n) + \nu \cdot I_{\text{SP}}(n). \tag{1.4}$$

Here, the expectation is with respect to the randomness in the channel conditions. Taking expectation with respect to the possible randomness of decision vector, applying law of iterated expectations followed by linearity of the expectation operator gives,

$$\begin{aligned} \mathbb{E}[d(n)] &= \mathbb{E}[\mathbb{E}[d(n) \mid \boldsymbol{I}(n)]] && (1.5) \\ &= \mathbb{E}\left[\mu \cdot I_{\text{SR}}(n) + \nu \cdot I_{\text{SP}}(n)\right] && (1.6) \\ &= \mu \cdot \mathbb{E}[I_{\text{SR}}(n)] + \nu \cdot \mathbb{E}[I_{\text{SP}}(n)] && (1.7) \\ &\triangleq \psi(n). && (1.8) \end{aligned}$$

**Note:-**

The variables $\mu$ and $\nu$ are a function of system-model parameters. These quantities, respectively, are the probability of success, averaged over the channel condition, of a raw packet and processed packet, respectively. Furthermore, by construction of $d(n)$, $\mathbb{E}[d(n)]$ denotes the probability of successful update, i.e. $P(d(n) = 1)$, in slot $n$. For convenience, we shall denote this quantity by $\psi(n)$. Thus,

$$P(d(n) = 1) \leftrightarrow \mathbb{E}[d(n)] \leftrightarrow \psi(n)$$

### 1.2.4 Cost of Actuation Error - A Context-Aware Distortion Measure

In many scenarios, including critical applications such as hazard detection, the destination performs some state-dependent action(s). It is essential that destination acquires updates from the source in timely and accurate manner (i.e., with low distortion). Decisions based on erroneous state estimates can lead to either unnecessary actuation or failure to act as and when required. Each of these outcomes is associated with a cost. For instance, unnecessary actuation would waste physical resources, while taking no action can lead to damage or system failure. This necessitates consideration of significance of the situation and context of the system into decision-making.

The Cost of Actuation Error (CAE) [Fountoulakis et al., 2023] quantifies the consequences and *semantics*, formally capturing the cost incurred due to mismatch between the true state and its estimate. Let $X(n)$ and $\hat{X}(n)$ denote the true state and the estimate at time $n$, respectively. Let instantaneous CAE be denoted by $\Delta(n)$.

**Definition 1.2.2: Cost of Actuation Error**

The CAE function, $\Delta : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is defined as

$$\Delta(n) = \Delta(X(n) = i, \hat{X}(n) = j) = \delta_{i,j}, \forall\, i, j \in \mathcal{X},$$

where $\delta_{i,j} \geqslant 0$ for $i \neq j$ and $\delta_{i,j} \leqslant 0$ otherwise.

The values $\delta_{i,j}$ encode the actuation costs when the estimated state is $j$ whereas the true state is $i$. Notably, the CAE is non-commutative, meaning that $\delta_{i,j} \neq \delta_{j,i}$ for $i \neq j$ in general. This asymmetry encapsulates the directional nature of decision costs. Consequently, the actuation error serves as a generalized, context-aware distortion measure. The long-term time-averaged expected CAE is given by

$$\bar{\Delta} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[\Delta(n)],$$

where the expectation is with respect to the channel conditions and action vector. We consider the following constraint:

$$\bar{\Delta} \leqslant d_0,$$

where $d_0 \in \mathbb{R}_{++}$ denotes the bound on the long-term time-averaged expected CAE.

Using rules of probability, it is possible to derive the explicit expression for the instantaneous expected CAE (see Theorem 5.1). A summary[5] of the results, the marginal and joint distributions of the true state and its estimate, is presented in table 1.1.

| | **Probabilities** | **Expression** |
|---|---|---|
| Marginal | $P_{X(n)}(0), P_{\hat{X}(n)}(0)$ | $\pi_0^X$ |
| | $P_{X(n)}(1), P_{\hat{X}(n)}(1)$ | $\pi_1^X$ |
| Joint | $P_{X(n),\hat{X}(n)}(0,0)$ | $(\pi_0^X)^2 + \pi_0^X\pi_1^X\psi(n)$ |
| | $P_{X(n),\hat{X}(n)}(0,1)$ | $\pi_0^X\pi_1^X(1-\psi(n))$ |
| | $P_{X(n),\hat{X}(n)}(1,0)$ | $\pi_0^X\pi_1^X(1-\psi(n))$ |
| | $P_{X(n),\hat{X}(n)}(1,1)$ | $(\pi_1^X)^2 + \pi_0^X\pi_1^X\psi(n)$ |

Table 1.1: Marginal and joint distributions of the true state $X(n)$ and its estimate $\hat{X}(n)$.

**Lemma 1.1** Expression for the expected instantaneous CAE

TThe instantaneous expected CAE is given by

$$\mathbb{E}[\Delta(n)] = \zeta + \xi\psi(n), \tag{1.9}$$

where $\zeta = \sum_{i,j} \delta_{i,j}\pi_i^X\pi_j^X$ and $\xi = \pi_0^X\pi_1^X \sum_{i,j}(-1)^{i+j}\delta_{i,j}$.

[5] See section 3.1 for a Monte-Carlo simulation based verification of these theoretical results.

*Proof.* This follows directly from using the results of Theorem 5.1 in calculating the expected value of $\Delta(n)$. Recollect that:

$$\Delta(n) = \delta_{i,j} \text{ if } X(n) = i \text{ and } \hat{X}(n) = j.$$

😄

The expected instantaneous CAE equation 1.9 comprises two components: $\zeta$ and $\xi\psi(n)$. We make the following remarks regarding their significance.

**Note:-**

1. The first term, $\zeta$, represents the expected CAE under the assumption that the true state and its estimate evolve independently, with identical marginal distributions. Under this assumption, the CAE contribution is $\delta_{i,j}$ whenever the true state is $X = i$ and the estimate is $\hat{X} = j$.

2. However, the estimate is not independent of the true state—particularly in the case of a successful update, where the estimate becomes a deterministic function of the actual state. Among the four possible state-estimate pairs—$(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$—a successful update corrects the mismatched pairs $(0,1)$ and $(1,0)$ with probability $\psi(n)$. Specifically, if a pair $(0,1)$ is corrected to $(0,0)$, the CAE must be reduced by $\delta_{0,1}$ and increased by $\delta_{0,0}$. Similarly, correction of a $(1,0)$ pair to $(1,1)$ leads to an adjustment of $\delta_{1,1} - \delta_{1,0}$. The second term, $\xi\psi(n)$, captures this correction to the expected CAE arising from the dependence of the estimate on the true state. It aggregates the expected net reduction in CAE across such events, weighted by their likelihood under the update mechanism.

This result plays a key role, albeit subtly, by enabling a straightforward derivation of the performance bound for the SRP, later on.

### 1.2.5 Age of Information Model

Let $A(n)$ denote the *Age of Information* at the destination at the beginning of time slot $n$. The evolution of $A(n)$ is governed by the following update rule:

$$A(n+1) = \begin{cases} w_0, & \text{if } \hat{X}(n) = 0, d(n) = 1, \\ w_1, & \text{if } \hat{X}(n) = 1, d(n) = 1, \\ A(n) + 1, & \text{otherwise.} \end{cases} \tag{1.10}$$

**Note:-**

The above given definition is a generalization of the state-specific AoI definition used in [S. et al., 2023], where one is concerned with state-aware age, i.e., the age of information only in a particular state. To get back this definition, set $w_0 = 0, w_1 = 1$. Many studies [?] use a broader, non-state-discriminatory definition of AoI. This definition can be recovered by setting $w_0 = w_1 = 1$.
To ensure that we do not miss anything by making this generalization, we should not allow there to be the expression '$(\frac{1}{w_1 - w_0})$' in any expression as it is invalid in case $w_1 = w_0$. This holds true allowing us to use this altered definition.
It would be interesting to find if another application-dependent interpretation can be form for this evolution form, since the math required is already developed here.

The *long-term time-average expected age*, denoted by $\bar{A}$, is the objective we seek to minimize while satisfying the previously discussed constraints. It is defined as

$$\bar{A} = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[A(n)]. \tag{1.11}$$

## 1.3 Optimization Analysis

### 1.3.1 The General Problem

We are now in a position to formally state the main optimization problem addressed in this paper. Let $\mathcal{P}$ denote the class of all non-anticipative policies, and let $\mathsf{P} \in \mathcal{P}$ represent an arbitrary admissible policy. The optimization problem is formulated as follows:

**Problem 1: General Problem**

$$\mathrm{OPT}^{\star} = \min_{\mathsf{P}\in\mathcal{P}} \quad \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[A(n)] \tag{1.12a}$$

$$\text{subject to} \quad \bar{\Delta} \leqslant d_0 \tag{1.12b}$$

$$\bar{C} \leqslant c_0 \tag{1.12c}$$

$$I_A(n)\cdot(I_A(n)-1) = 1, \quad \forall\, A \in \mathcal{A},\; n = 1, 2, \ldots \tag{1.12d}$$

$$\sum_{A\in\mathcal{A}} I_A(n) = 1, \quad \forall\, n = 1, 2, \ldots \tag{1.12e}$$

where $\bar{\Delta}$ and $\bar{C}$ are as defined previously.

In the rest of the text, we refer to the equations 1.12a-1.12e as the *general problem*.

### 1.3.2 Stationary Randomized Policies (SRPs)

Let $\mathcal{P}_R$ denote the set of all SRPs which choose an action with a certain probability in every slot. Concretely, a policy $\mathsf{R} \in \mathcal{P}_R$ selects an action $A \in \mathcal{A}$ with probability $p_A^{\mathsf{R}}$ in slot $n$. That is, the policy $\mathsf{R}$ is fully characterized by the probability vector $\boldsymbol{p}_{\mathsf{R}} = [p_{\mathrm{NS}}^{\mathsf{R}}, p_{\mathrm{SR}}^{\mathsf{R}}, p_{\mathrm{SP}}^{\mathsf{R}}]$. In the rest of this section, we reformulate the primary general problem under the class of SRPs and present a reformulated convex optimization problem to obtain $\boldsymbol{p}_{\mathsf{R}}$. We then demonstrate that the optimal AoI objective achieved under this optimal SRP is within a bounded multiplicative gap from that of the original general problem.

#### Reformulation

Under SRP $\mathsf{R}$ we have, $\mathbb{E}[I_A(n)] = p_A^{\mathsf{R}}, \forall\, A \in \mathcal{A}, n \in \mathbb{N}$. Substituting this into equation 1.8 yields:

$$\psi^{\mathsf{R}}(n) = \psi^{\mathsf{R}} = \mu p_{\mathrm{SR}}^{\mathsf{R}} + \nu p_{\mathrm{SP}}^{\mathsf{R}}, \tag{1.13}$$

and we can also show the following:

**Theorem 1.1** Expression for long-term time-averaged expected AoI

Let $\mathsf{R} \in \mathcal{P}_R$ be an arbitrary SRP. Under this policy, the long-term time-averaged expected AoI is given by:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[A(n)] = \left(\frac{1}{\psi^{\mathsf{R}}} - 1\right) + w_0 + \pi_1^X(w_1 - w_0). \tag{1.14}$$

*Proof.* Refer Theorem 5.3 for a detailed proof. 😄

We can now put together the problem that one must solve in order to find the optimal SRP:

**Problem 2: SRP Problem**

$$
\begin{aligned}
\mathrm{OPT}^{\star}_{\mathsf{R}} = \min_{\mathsf{R}\in\mathcal{P}_R} \quad & \left(\frac{1}{\psi^{\mathsf{R}}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0) && (1.15a)\\
\text{subject to} \quad & \zeta + \xi\psi^{\mathsf{R}} \leqslant d_0, && (1.15b)\\
& \boldsymbol{c}^\top \boldsymbol{p} \leqslant c_0, && (1.15c)\\
& p_A^{\mathsf{R}} \geqslant 0, \quad \forall\, A \in \mathcal{A}. && (1.15d)\\
& \sum_{A\in\mathcal{A}} p_A^{\mathsf{R}} = 1. && (1.15e)
\end{aligned}
$$

We refer to the problem equation 1.15a - equation 1.15e as the *SRP problem*. The constraints in equation 1.15b and equation 1.15b follow from the time-invariant and stochastic nature of decision-making in SRPs. Additionally, the conditions in equation 1.15d and equation 1.15e ensure that the SRP adheres to the validity constraints of the general problem. It can be verified that the above problem is convex. Consequently, it can be efficiently solved using standard convex optimization solvers.

### Performance Bound

Policies in this class may be suboptimal. Therefore, we now establish a performance bound for this class of policies. To achieve this, we consider an alternative problem that serves as a lower bound for the general problem.

**Problem 3: Lower Bound Problem**

$$
\begin{aligned}
L_B^{\star} = \min_{\mathsf{p}\in\mathcal{P}} \quad & \frac{1}{2}\left(\frac{1}{\hat{q}} - 1\right) + w_0 + \pi_1^X (w_1 - w_0) && (1.16a)\\
\text{subject to} \quad & \bar{\Delta} \leqslant d_0, && (1.16b)\\
& \bar{C} \leqslant c_0, && (1.16c)\\
& I_A(n)(I_A(n) - 1) = 1, \forall\, A \in \mathcal{A},\ n \in \mathbb{N}, && (1.16d)\\
& \sum_{A\in\mathcal{A}} I_A(n) = 1, \forall\, n \in \mathbb{N}, && (1.16e)
\end{aligned}
$$

where $\hat{q}$ denotes the long-term throughput of the source and defined as:

$$
\hat{q} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[d(n)]. \tag{1.17}
$$

The problem equation 1.16a - equation 1.16e will be referred to as the *lower-bound problem.*

**Lemma 1.2** Lower Bound Problem provides a lower bound to the General Problem

The solution of the lower bound problem is a lower bound of the solution to the general problem. Mathematically, we have

$$
L_B^{\star} \leqslant \mathrm{OPT}^{\star}, \tag{1.18}
$$

for every instance of our system model.

*Proof.* Following an approach similar to that in [Kadota et al., 2018], we establish that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[A(n)] \geqslant \frac{1}{2}\left(\frac{1}{\hat{q}} - 1\right) + w_0 + \pi_1^X (w_1 - w_0).$$

Due to the structural differences in the AoI evolution model, we employ a modified *partially-expected sample path* approach rather than the single sample path approach used in [Kadota et al., 2018]. Substituting this lower bound into the objective function shows that the objective value of the original problem (equation 1.12a) is lower bounded by that of the lower bound problem (equation 1.16a). 😄

Taking advantage of the similarity between the objective functions in equation 1.15a and equation 1.16a, we derive a performance bound, which is formally stated in the following theorem.

**Theorem 1.2** Optimality Bound on SRP class solutions.
For every instance of our system model, we have

$$\mathrm{OPT}^\star_{\mathsf{R}} \leqslant 2\mathrm{OPT}^\star.$$

*Proof.* Let $\mathsf{P}_{L_B}$ denote the policy that solves the problem of the lower bound. Let $\bar{\Delta}_{\mathsf{P}_{L_B}}$ and $\bar{C}_{\mathsf{P}_{L_B}}$ be the long-term time average of the expected CAE and cost of actions associated with this policy. Construct a stationary randomized policy $\mathsf{R}_0 \in \mathcal{P}_R$ such that,

$$p^{\mathsf{R}_0}_{\mathrm{NS}} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[I_{\mathrm{NS}}(n)]_{\mathsf{P}_{L_B}}, \tag{1.19a}$$

$$p^{\mathsf{R}_0}_{\mathrm{SR}} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[I_{\mathrm{SR}}(n)]_{\mathsf{P}_{L_B}}, \tag{1.19b}$$

and

$$p^{\mathsf{R}_0}_{\mathrm{SP}} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[I_{\mathrm{SP}}(n)]_{\mathsf{P}_{L_B}}. \tag{1.19c}$$

By exploiting the linearity of expectation, limits and summation, it is a straightforward[6] exercise to show that $\bar{\Delta}^{\mathsf{P}_{L_B}} = \bar{\Delta}^{\mathsf{R}_0}$ and $\bar{C}^{\mathsf{P}_{L_B}} = \bar{C}^{\mathsf{R}_0}$. Thus, $\mathsf{R}_0$ satisfies all constraints if $\mathsf{P}_{L_B}$ is feasible. We conclude that $\mathsf{R}_0$ is feasible if $\mathsf{P}_{L_B}$ is feasible. Furthermore, by using the linearity of expectation, limits, and summation again, we can show that $\hat{q}^{\mathsf{P}_{L_B}} = \psi^{\mathsf{R}_0}$.

Comparing the objectives,

$$\frac{\mathrm{OPT}_{\mathsf{R}_0}}{2} = \frac{1}{2}\left(\frac{1}{\psi^{\mathsf{R}_0}} - 1\right) + \frac{w_0 + \pi_1^X (w_1 - w_0)}{2} \tag{1.20}$$

$$< \frac{1}{2}\left(\frac{1}{\psi^{\mathsf{R}_0}} - 1\right) + w_0 + \pi_1^X (w_1 - w_0) \tag{1.21}$$

$$= \frac{1}{2}\left(\frac{1}{\hat{q}^{\mathsf{P}_{L_B}}} - 1\right) + w_0 + \pi_1^X (w_1 - w_0) \tag{1.22}$$

$$= L_B^\star. \tag{1.23}$$

Thus, $\mathrm{OPT}_{\mathsf{R}_0} < 2L_B^\star$. Recall that from lemma 1.2 $L_B^\star \leqslant \mathrm{OPT}^\star \leqslant \mathrm{OPT}^\star_{\mathsf{R}} \leqslant \mathrm{OPT}_{\mathsf{R}_0}$. Hence,

$$\frac{\mathrm{OPT}^\star_{\mathsf{R}}}{\mathrm{OPT}^\star} \leqslant \frac{\mathrm{OPT}_{\mathsf{R}_0}}{\mathrm{OPT}^\star} \leqslant \frac{\mathrm{OPT}_{\mathsf{R}_0}}{L_B^\star} < 2. \tag{1.24}$$

😄

[6] In this step, we exploit the the closed-form expression for $\mathbb{E}[\Delta(n)]$ derived in Lemma 1.1.

**Note:-**

In the last theorem, we proved that $\mathrm{OPT}^\star_{\mathsf{R}}/\mathrm{OPT}^\star < 2$. Referring to the definition of optimality ratio given in [Kadota et al., 2018], we conclude that the SRP class of policies is *2-optimal*.

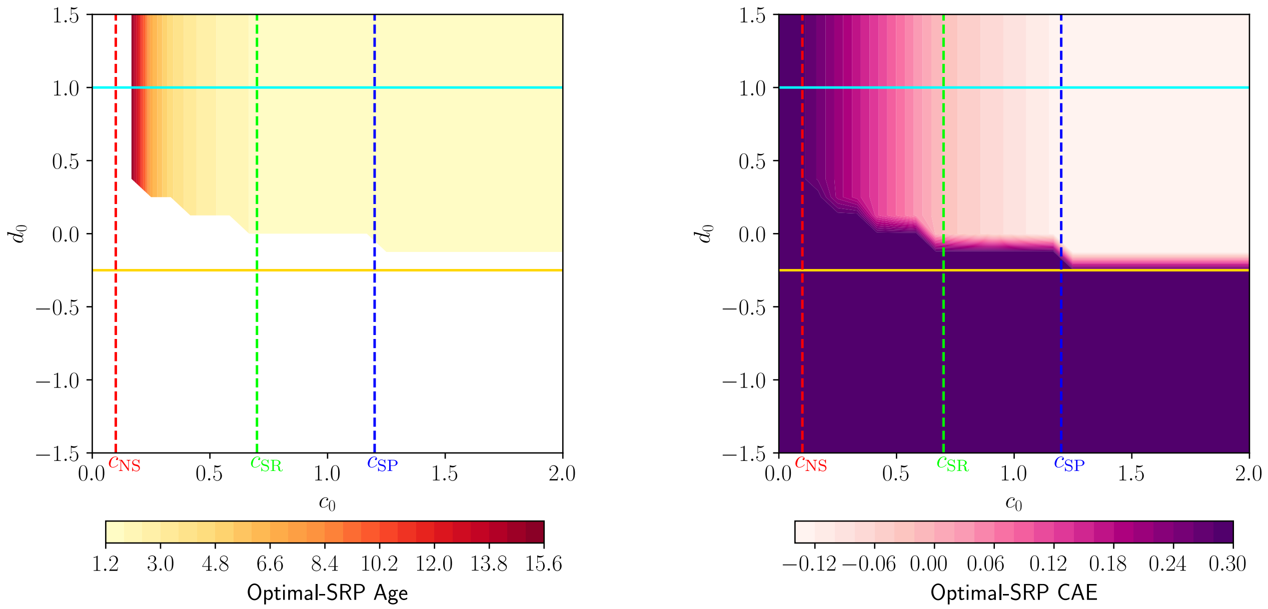
This concludes the mathematical analysis of the problem from the SRP class perspective. We restricted our study to this class of policies as it provides a solution that is trivial to implement. This demonstrates the trade-off between the performance of the policy and convenience of implementing it/operating overheads.

# Chapter 2

# Numerical Analysis

In this section, we numerically evaluate the impact of various system parameters on AoI and CAE under the SRP, including an assessment of problem feasibility. We further examine how these parameters influence the optimal action selection probabilities. Finally, we study the trade-off between AoI and CAE under the SRP. Unless mentioned otherwise, the model parameters were held constant as listed in table 2.1.

### 2.0.1 Feasibility Analysis




(a) (b)

Figure 2.1: Variation of long-term averaged metrics under the optimal SRP with respect to constraint upper bounds: (figure 2.1a) AoI, and (figure 2.1b) CAE.

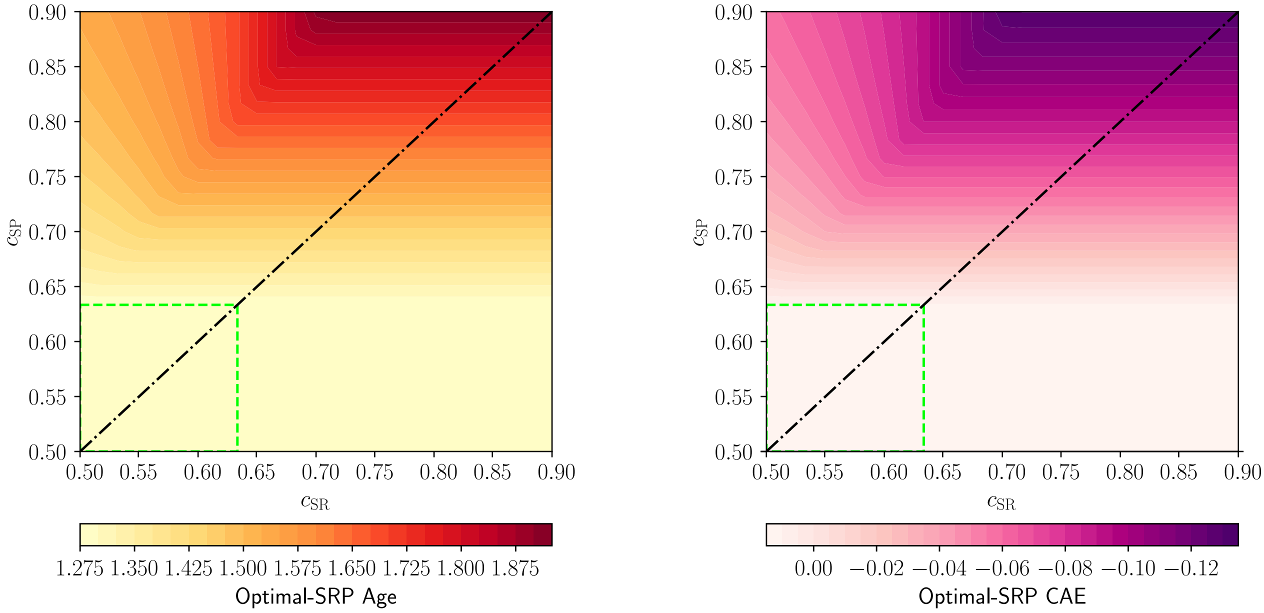
We study the set of constraint bounds $(c_0, d_0)$ for which the SRP optimization problem admits a solution, given a fixed set of parameters. Specifically, we vary $c_0$ over $[0.0, 2.0]$ and $d_0$ over $[-1.5, 1.5]$, while holding other parameters constant. The resulting optimal values of AoI and CAE are shown in figure 2.1a and figure 2.1b, respectively. Ideally, the feasible region for the AoI should be convex. However, due to the discretization of $c_0$ and $d_0$ values, the lower boundary of the region exhibits fluctuations.

For interpretability, we partition each plot into three vertical segments—Regions A, B, and C—based on cost thresholds: Region A spans $c_0 \in [c_{NS}, c_{SR})$, Region B spans $[c_{SR}, c_{SP})$, and Region C corresponds to $c_0 \geqslant c_{SP}$.

In Region A, where $c_0$ is close to $c_{NS}$, the tight cost constraint limits sampling frequency, resulting in high AoI and CAE. Here, feasibility is primarily governed by the CAE constraint. For small $d_0$ values, the problem becomes infeasible as the system cannot meet strict CAE requirements under a limited budget.

Table 2.1: Parameter values used for numerical analysis.

| **Parameter** | **Value** |
|---|---|
| $p^X_{0,1}$, $p^X_{1,0}$ | 0.35, 0.75 |
| $p_{\text{chnl}}$ | 0.6 |
| $p^p_1$, $p^p_0$, $p^r_1$, $p^r_0$ | 0.9, 0.6, 0.75, 0.35 |
| $\delta_{0,0}$, $\delta_{0,1}$, $\delta_{1,0}$, $\delta_{1,1}$ | $-0.25$, 1, 1, $-0.25$ |
| $c_0$, $d_0$ | 0.63, 0.2 |
| $c_{\text{NS}}, c_{\text{SR}}, c_{\text{SP}}$ | 0.1, 0.7, 1.2 |
| $w_0$, $w_1$ | 1, 1 |




(a) (b)

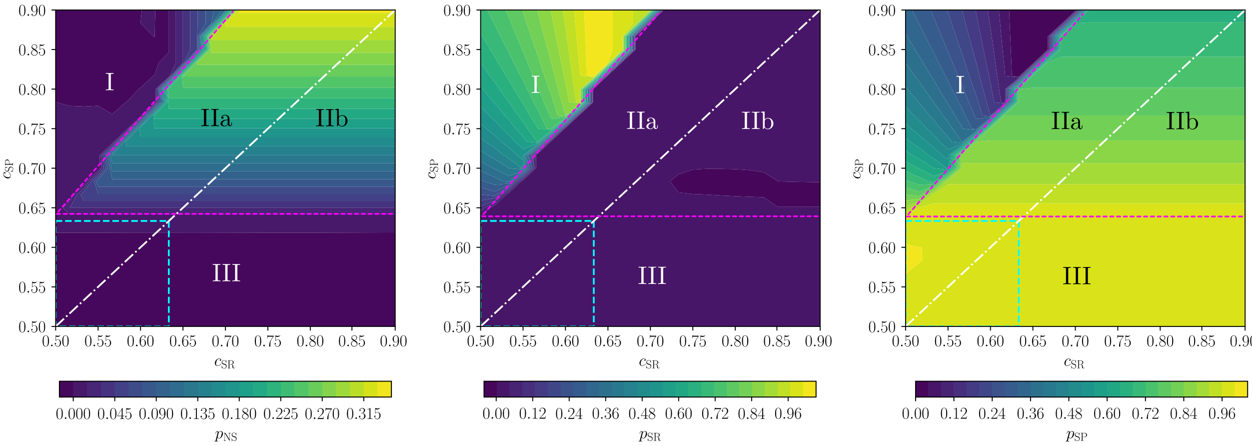
Figure 2.2: Impact of varying action costs on the observed long-term time-averaged metrics. The plot (figure 2.2a) illustrates the objective value, representing the long-term time-averaged AoI under the optimal SRP policy. The plot (figure 2.2b) shows the corresponding constraint value, depicting the CAE under the same policy.

Increasing $d_0$ relaxes this constraint, restoring feasibility. However, AoI remains high due to persistent sampling limitations. As $c_0$ increases beyond $c_{\text{SR}}$, the system can sample more frequently, perhaps first enabling raw data transmission (Region B), and eventually processed data (Region C) once $c_0 \geq c_{\text{SP}}$. The latter benefits from higher delivery success due to reduced packet sizes.

Notably, in the AoI plot, most contours are nearly horizontal—i.e., AoI remains largely unaffected by changes in $d_0$ for a fixed $c_0$. This confirms that AoI is primarily influenced by the cost bound under the current setting. The CAE bound $d_0$ has a limited impact on AoI, except near the infeasibility boundary where overly strict constraints render the problem unsolvable. Both the plots show a sharp transition in feasibility as $d_0$ increases from negative values toward zero. This indicates that even slight relaxations in the constraints can significantly expand the feasible region, particularly under low-cost budgets.

### 2.0.2 Impact of Action Costs

We now analyze how variations in transmission costs affect the system's long-term performance and optimal behavior. Specifically, we fix $c_{\text{NS}} = 0.1$ and vary $c_{\text{SR}}$ and $c_{\text{SP}}$ over the range $[0.5, 0.9]$, keeping the total cost constraint $c_0$ and other parameters fixed.




(a) (b) (c)

Figure 2.3: Effect of sampling costs on the optimal SRP policy. The optimal policy is characterized by the probabilities of different actions which are plotted as: (figure 2.3a) the probability of not sampling, (figure 2.3b) the probability of sampling and transmitting raw data, and (figure 2.3c) the probability of sampling and transmitting processed data (the state of the process).

Figure 2.2a and figure 2.2b present the observed long-term time-averaged AoI and CAE, respectively, under the optimal SRP. The corresponding policy consisting of the action probabilities $p_{\mathrm{NS}}, p_{\mathrm{SR}}$ and $p_{\mathrm{SP}}$ are shown in figures 2.3a, 2.3b, and 2.3c, respectively. For ease of interpretation, all plots are segmented into three distinct and major regions (I, II, and III), each reflecting qualitatively different system behavior.

Each plot includes the following annotations: (i) a green/cyan dashed box indicating the region bounded by $c_0$ on each axis, (ii) a black/white dash-dotted line corresponding to $c_{\mathrm{SR}} = c_{\mathrm{SP}}$, and (iii) pink dotted lines demarcating the boundaries between the three identified regions.

In Region III, both AoI and CAE are significantly lower than in other regions. As illustrated in Fig. 2.3, this can be attributed to the feasibility of executing the action SP in almost every time slot, given that $c_{\mathrm{SP}} \leqslant c_0$. Additionally, the higher success probability associated with processed packets ($\nu$) reduces both AoI and CAE. Consequently, the optimal SRP in this region predominantly favors the SP action.

Region II can be further subdivided into subregions IIa (above the line $c_{\mathrm{SR}} = c_{\mathrm{SP}}$) and IIb (below the line $c_{\mathrm{SR}} = c_{\mathrm{SP}}$). In Region IIb, where $c_{\mathrm{SR}} \geqslant c_{\mathrm{SP}}$, processing before transmission is more cost-effective and offers higher success probability. However, since $c_{\mathrm{SP}} > c_0$, the cost constraint forces a reduction in sampling frequency, thereby increasing the probability of the no-sampling action (NS). This results in increased AoI and CAE. In Region IIa, although $c_{\mathrm{SP}} > c_{\mathrm{SR}}$ suggests a preference for the SR action, Fig. 2.3 shows a continued dominance of the SP action. This apparent contradiction highlights the role of differing success probabilities $\mu$ (for raw data) and $\nu$ (for processed data). A substantially higher $\nu$ can outweigh the cost advantage of the SR action, especially if $\mu$ is relatively low. As a result, the optimal SRP still favors SP, alongside a relatively high $p_{\mathrm{NS}}$ compared to Region IIb.

In Region I, where the cost asymmetry between raw and processed data transmission is most pronounced, the optimal SRP predominantly employs the SR action, occasionally interspersed with SP (close to $c_0$) and very rarely chooses NS. This behavior is expected, as the system must compensate for the high cost of SP while still meeting the constraint. Two strategies are available: (i) transmit raw samples frequently and use processed data transmissions sparingly, or (ii) reduce sampling frequency. Since $c_{\mathrm{SR}} \leqslant c_0$ in this region, the former strategy becomes feasible. As a result, the optimal SRP in Region I delivers intermediate performance in terms of both AoI and CAE.

### 2.0.3 Tradeoffs Between AoI and CAE

Figure 2.4 illustrates the tradeoff between the average AoI and average CAE achieved by the optimal SRP under various system configurations. This analysis aims to understand how variations in the source parameters (transition probabilities) and channel reliability influence these performance metrics. Although

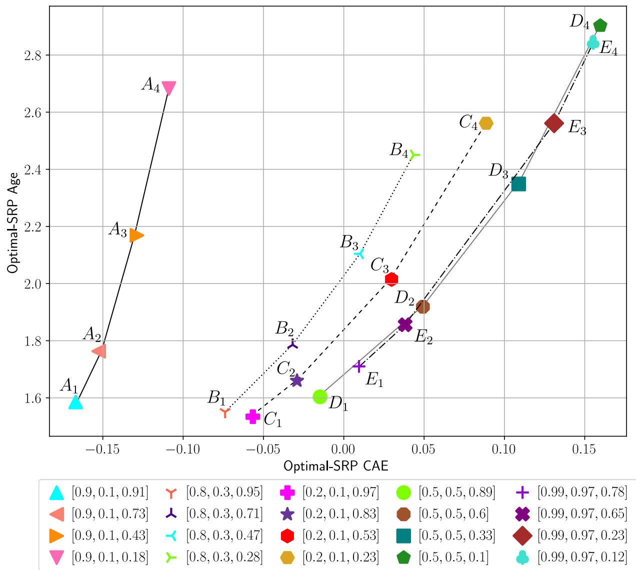



Figure 2.4: The average Age of Information (AoI) and average CAE observed under the optimal SRP policy across various instances of the system model. Each legend entry denotes the tuple $(p^X_{0,1}, p^h_{1,0}, p_{\text{chnl}})$, corresponding to the parameters governing the exogenous process, hidden process, and channel reliability, respectively.

figure 2.2 suggests that AoI and CAE often increase or decrease together—potentially implying that minimizing one also minimizes the other—this is not always the case, as highlighted in figure 2.4.

Consider the trajectories $A_1 \to A_4$, $B_1 \to B_4$, and $C_1 \to C_4$ in figure 2.4, which correspond to decreasing channel success probability $p_{\text{chnl}}$. When the source transition probabilities are fixed (i.e., fixed $(p^X_{0,1}, p^X_{1,0})$), both AoI and CAE increase monotonically as channel reliability degrades. This is intuitive—fewer successful transmissions lead to staler information and less accurate estimates.

However, consider points $A_2$, $B_2$, and $E_2$, which have nearly identical $p_{\text{chnl}}$. These points exhibit similar AoI but significantly different CAE levels. Since the channel reliability is constant, the variation in CAE arises from differences in the source transition probabilities. Specifically, when the source transitions are asymmetric (e.g., one high and one low), the system tends to dwell longer in the state with the lower transition probability. Because the estimate is typically held constant between updates, the likelihood of matching the actual state is higher, resulting in lower CAE. In contrast, when both transition probabilities are high (e.g., $(p^X_{0,1}, p^X_{1,0}) = (0.5, 0.5)$ or $(0.9, 0.9)$), the state changes frequently, increasing the risk of mismatch and hence CAE.

In summary, this figure emphasizes that AoI and CAE capture fundamentally different dimensions of system performance. Optimizing one may not guarantee optimality in the other. For systems where both information freshness and accuracy matter, sampling and transmission strategies must be designed with joint consideration of both metrics.

# Chapter 3

# Simulation Methodology

## 3.1 Information Distortion

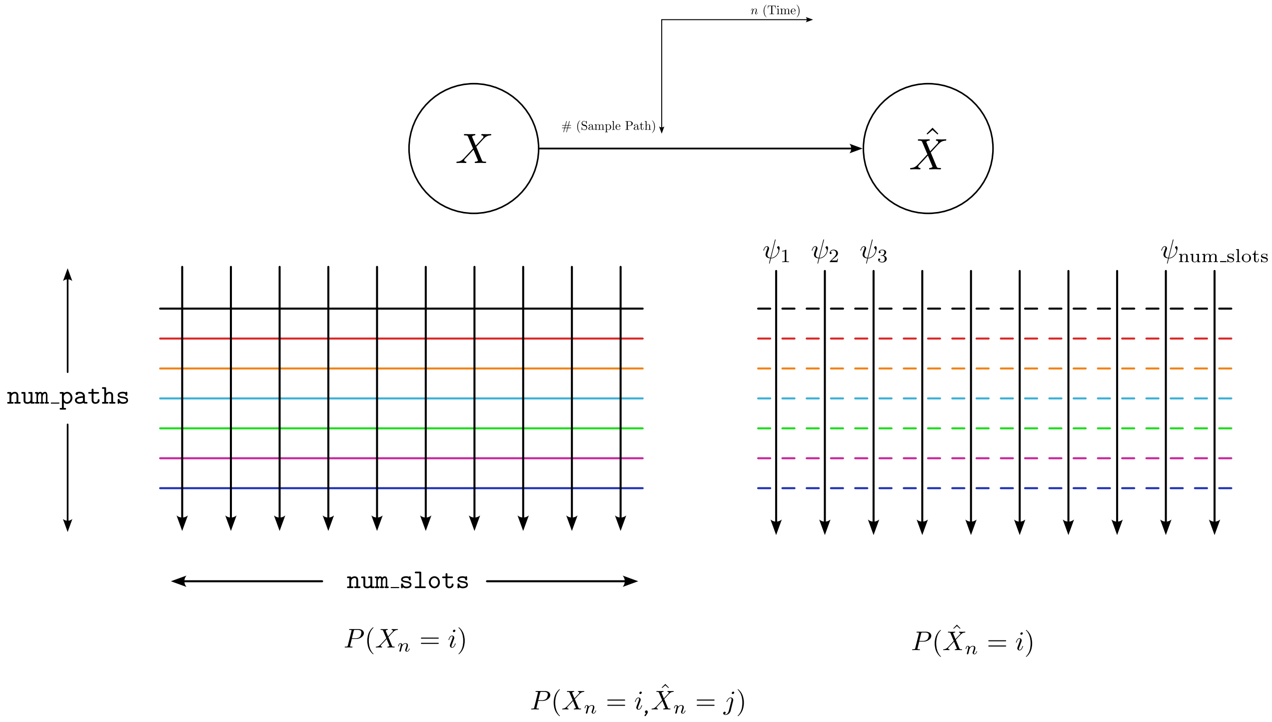



Figure 3.1: Pictorial representation of the simulations conducted to numerically obtain the marginal and joint distributions of the true Markov chain and its sampled observations.

### 3.1.1 Simulation Methodology

To validate the theoretical expressions for the marginal and joint distributions of the true state and its estimate, we perform numerical simulations. The theoretical distributions are defined as follows:

$$\pi^X = \begin{bmatrix} [P_{X(1)}(0) & P_{X(1)}(1)] \\ [P_{X(2)}(0) & P_{X(2)}(1)] \\ \vdots & \\ [P_{X(N)}(0) & P_{X(N)}(1)] \end{bmatrix}, \tag{3.1}$$

$$\pi^{\hat{X}} = \begin{bmatrix} [P_{\hat{X}(1)}(0) & P_{\hat{X}(1)}(1)] \\ [P_{\hat{X}(2)}(0) & P_{\hat{X}(2)}(1)] \\ \vdots & \\ [P_{\hat{X}(N)}(0) & P_{\hat{X}(N)}(1)] \end{bmatrix}, \tag{3.2}$$

and the joint distributions:

$$P_{X(n),\hat{X}(n)} = \begin{bmatrix} [P_{X(1),\hat{X}(1)}(0,0) & P_{X(1),\hat{X}(1)}(0,1) & P_{X(1),\hat{X}(1)}(1,0) & P_{X(1),\hat{X}(1)}(1,1)] \\ [P_{X(2),\hat{X}(2)}(0,0) & P_{X(2),\hat{X}(2)}(0,1) & P_{X(2),\hat{X}(2)}(1,0) & P_{X(2),\hat{X}(2)}(1,1)] \\ \vdots & & & \\ [P_{X(N),\hat{X}(N)}(0,0) & P_{X(N),\hat{X}(N)}(0,1) & P_{X(N),\hat{X}(N)}(1,0) & P_{X(N),\hat{X}(N)}(1,1)] \end{bmatrix}. \tag{3.3}$$

We generate uniformly random $\psi(n)$ values from $(0, 1)$ for all $n$, use them to compute the above distributions, and conduct simulations based on Algorithm 1. The empirical distributions, obtained from simulated sample paths, are defined as:

$$\tilde{\pi}^X = \begin{bmatrix} [\tilde{P}_{X(1)}(0) & \tilde{P}_{X(1)}(1)] \\ [\tilde{P}_{X(2)}(0) & \tilde{P}_{X(2)}(1)] \\ \vdots & \\ [\tilde{P}_{X(N)}(0) & \tilde{P}_{X(N)}(1)] \end{bmatrix}, \tag{3.4}$$

$$\tilde{\pi}^{\hat{X}} = \begin{bmatrix} [\tilde{P}_{\hat{X}(1)}(0) & \tilde{P}_{\hat{X}(1)}(1)] \\ [\tilde{P}_{\hat{X}(2)}(0) & \tilde{P}_{\hat{X}(2)}(1)] \\ \vdots & \\ [\tilde{P}_{\hat{X}(N)}(0) & \tilde{P}_{\hat{X}(N)}(1)] \end{bmatrix}, \tag{3.5}$$

and the joint distributions:

$$\tilde{P}_{X(n),\hat{X}(n)} = \begin{bmatrix} [\tilde{P}_{X(1),\hat{X}(1)}(0,0) & \tilde{P}_{X(1),\hat{X}(1)}(0,1) & \tilde{P}_{X(1),\hat{X}(1)}(1,0) & \tilde{P}_{X(1),\hat{X}(1)}(1,1)] \\ [\tilde{P}_{X(2),\hat{X}(2)}(0,0) & \tilde{P}_{X(2),\hat{X}(2)}(0,1) & \tilde{P}_{X(2),\hat{X}(2)}(1,0) & \tilde{P}_{X(2),\hat{X}(2)}(1,1)] \\ \vdots & & & \\ [\tilde{P}_{X(N),\hat{X}(N)}(0,0) & \tilde{P}_{X(N),\hat{X}(N)}(0,1) & \tilde{P}_{X(N),\hat{X}(N)}(1,0) & \tilde{P}_{X(N),\hat{X}(N)}(1,1)] \end{bmatrix}. \tag{3.6}$$

### 3.1.2 Closeness Metrics

To quantify the similarity between the theoretical and empirical distributions, we use two statistical divergence measures:

**Definition 3.1.1: Symmetric Kullback-Leibler (sKL) Divergence**

The Kullback-Leibler (KL) distance between probability mass functions $P$ and $Q$ is given by:

$$KL(P \parallel Q) = \sum_{x \in \mathcal{X}} p(x) \log \left( \frac{p(x)}{q(x)} \right), \tag{3.7}$$

where $\mathcal{X}$ is the support of $P$ and $Q$. The symmetric KL-divergence, also known as just divergence, is defined as:

$$D_K(P, Q) = KL(P \parallel Q) + KL(Q \parallel P). \tag{3.8}$$

**Definition 3.1.2: Bhattacharyya Distance**

The Bhattacharyya distance is given by:

$$D_B(P, Q) = -\ln\left(\mathrm{BC}(P, Q)\right), \tag{3.9}$$

where

$$\mathrm{BC}(P, Q) = \sum_{x \in \mathcal{X}} \sqrt{p(x)q(x)} \tag{3.10}$$

is the Bhattacharyya coefficient of $P$ and $Q$. This metric provides an alternative measure of closeness.

We calculate the above distances for every time slot resulting in the following data:

$$\mathrm{KL}_X = \left[D_K(P_{X(1)}, \tilde{P}_{X(1)}), \ldots, D_K(P_{X(N)}, \tilde{P}_{X(N)})\right], \tag{3.11}$$

$$\mathrm{KL}_{\hat{X}} = \left[D_K(P_{\hat{X}(1)}, \tilde{P}_{\hat{X}(1)}), \ldots, D_K(P_{\hat{X}(N)}, \tilde{P}_{\hat{X}(N)})\right], \tag{3.12}$$

$$\mathrm{KL}_{X,\hat{X}} = \left[D_K(P_{X(1),\hat{X}(1)}, \tilde{P}_{X(1),\hat{X}(1)}), \ldots, D_K(P_{X(N),\hat{X}(N)}, \tilde{P}_{X(N),\hat{X}(N)})\right], \tag{3.13}$$

$$\mathrm{BD}_X = \left[D_B(P_{X(1)}, \tilde{P}_{X(1)}), \ldots, D_B(P_{X(N)}, \tilde{P}_{X(N)})\right], \tag{3.14}$$

$$\mathrm{BD}_{\hat{X}} = \left[D_B(P_{\hat{X}(1)}, \tilde{P}_{\hat{X}(1)}), \ldots, D_B(P_{\hat{X}(N)}, \tilde{P}_{\hat{X}(N)})\right], \tag{3.15}$$

$$\tag{3.16}$$

and

$$\mathrm{BD}_{X,\hat{X}} = \left[D_B(P_{X(1),\hat{X}(1)}, \tilde{P}_{X(1),\hat{X}(1)}), \ldots, D_B(P_{X(N),\hat{X}(N)}, \tilde{P}_{X(N),\hat{X}(N)})\right]. \tag{3.17}$$

## 3.2 Results and Analysis

The computed divergence values over time are presented in figure figure 3.2, while the corresponding mean and variance are summarized in table 3.1. Additionally, figure figure 3.3 provides a visual representation of the simulation analysis. The observed metric values were found to indicate a strong agreement between the theoretical (predicted) and empirical (observed) distributions.

| Metric | Distribution | Mean | Variance |
|---|---|---|---|
| sKL-Divergence | $P_{X(n)}, \tilde{P}_{X(n)}$ | 9.69e-05 | 1.84e-08 |
| | $P_{\hat{X}(n)}, \tilde{P}_{X(n)}$ | 0.00011 | 2.29e-08 |
| | $P_{X(n),\hat{X}(n)}$ | 0.0139 | 0.000455 |
| Bhattacharyya Distance | $P_{X(n)}, \tilde{P}_{X(n)}$ | 1.21e-05 | 2.87e-10 |
| | $P_{\hat{X}(n)}, \tilde{P}_{X(n)}$ | 1.38e-05 | 3.58e-10 |
| | $P_{X(n),\hat{X}(n)}$ | 0.0017 | 7.16e-06 |

Table 3.1: Summary of statistical divergence measures comparing theoretical and empirical distributions.

**Algorithm 1:** Simulation of a Discrete-Time Markov Chain with random sampling and computation of divergence measures for marginal and joint state distributions.

```
Input: num_slots, num_paths, stateTransMat, psi
Output: Empirical distributions and statistical divergence metrics
for i = 1 to num_paths do
    Sample initial state from X distribution
    for n = 1 to num_slots do
        Sample next state using stateTransMat
    Save the sample path in X_paths
for i = 1 to num_paths do
    Sample each path from X_paths estimated state based on ψ(n)
    Save these "sampled" sample path in X_hat_paths
Compute theoretical and empirical distributions
Compute sKL-divergence and Bhattacharyya distance
```

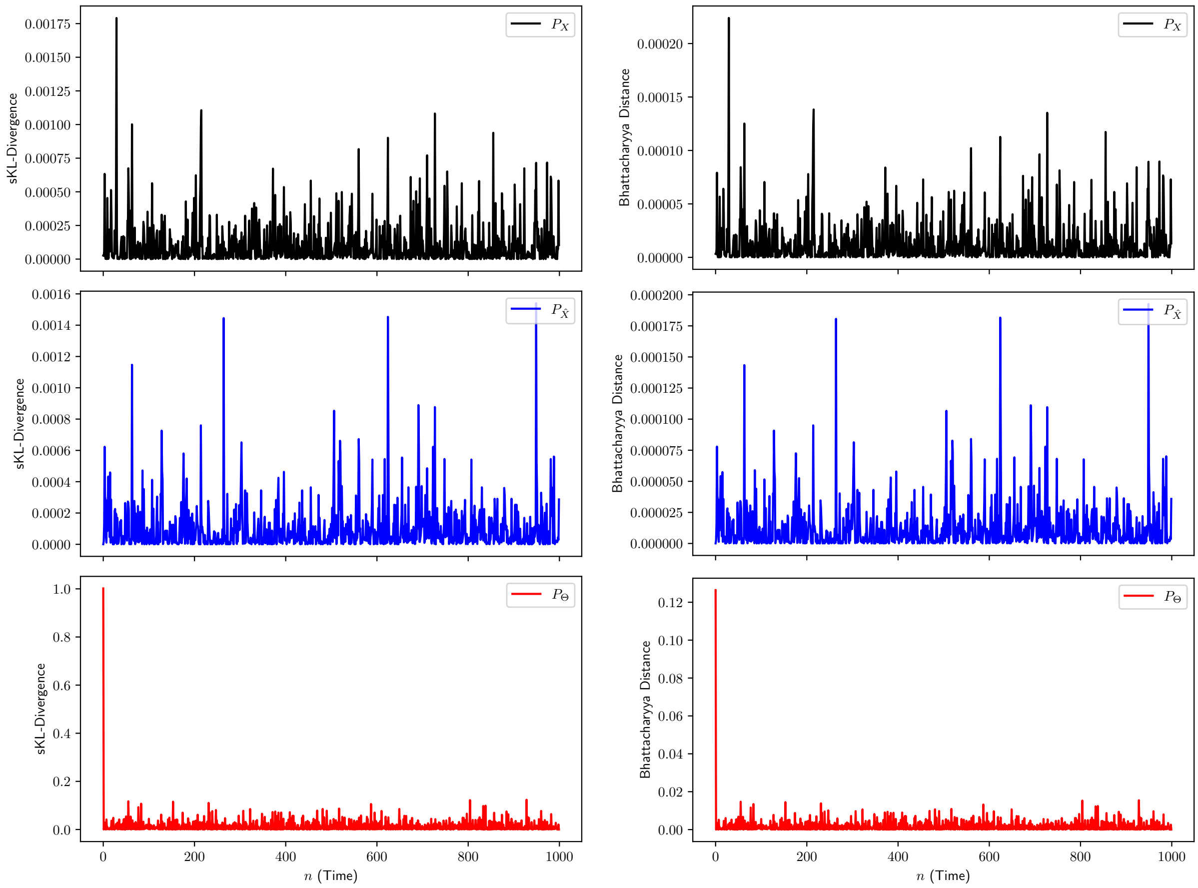



Figure 3.2: Evolution of statistical divergence measures over time, averaged over 10,000 sample paths.

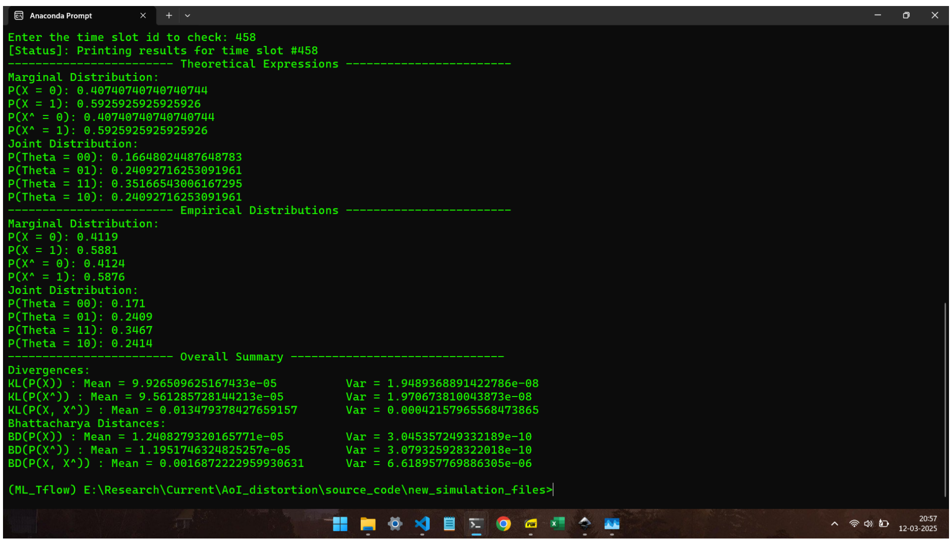



Figure 3.3: Screenshot of the post-simulation analysis.

# Chapter 4

# A Swarm of Sensors

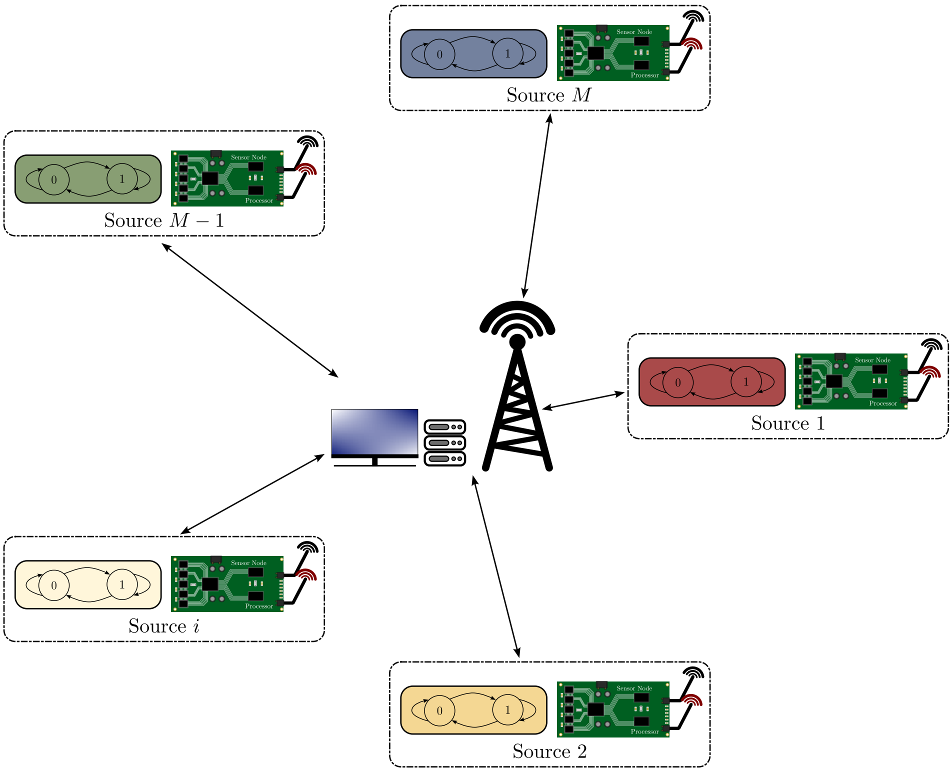



Figure 4.1: We extend the previously developed theory to a network of sensor nodes, a multiple-source single-destination set-up. We impose an interference constraint that at any given time utmost one sensor node may transmit.

Having examined the single-sensor setup, we now extend our analysis to a networked multi-sensor scenario. (see figure 4.1). In this, our ultimate goal is to monitor heterogeneous (non-identical) processes. Such a generalization requires substantial rethinking and deeper analysis. For tractability and consistency, we introduce a set of simplifying assumptions that preserve allow the use of previously proven results.

We consider a single-hop wireless network in which a controller (destination) receives updates from $M$ sensor (source) nodes. Time is slotted, indexed by $n = 1, 2, \ldots$, and the wireless channel supports at most one transmission per slot. In each slot n, the destination either remains idle or schedules one node $m \in \mathcal{M} = \{1, \ldots, M\}$ to transmit. Upon selection, a node can transmit either a raw or a processed data packet. A schematic of a single node's connection to the destination is provided in figure 1.3.

The subsequent sections detail the models for the source processes, communication channel, constraints on the Cost of Actuation Error (CAE) and transmission costs, and the Age of Information (AoI) metric.

## 4.1 System Model

### 4.1.1 Sources

We will model the exogenous two-state process as a Discrete Time Markov Chain (DTMC). Each node may monitor a process with different statistics. Let the DTMC corresponding to the $m^{\text{th}}$ node have the states denoted by $X_m(1), X_m(2), \ldots$. Let $\mathcal{X}_m = \{0_m, 1_m\}$[1] be the state-space and $\boldsymbol{\pi}_{X_m} = [\pi^X_{0,m}, \pi^X_{1,m}]$ be the stationary distribution of the $m^{th}$ process. The probability of the $m^{\text{th}}$ process transitioning from state $i$ to state $j$ (where $i, j \in \mathcal{X}$) is given by $p^{X_m}_{i,j} \in [0, 1]$ for all $n$.

### 4.1.2 Channel

The sensors send their data over a single-hop wireless channel. The channel itself permits a single transmission per time slot. Further, the channel assumes two states - bad (0) and good (1). If $H(n)$ denotes the channel condition in slot $n$, $H(n) \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p_{\text{chnl}}),$ where $p_{\text{chnl}} \in [0, 1]$.

Let $p^t_{h,m}$ denote the probability of a packet transmitted by node $m$, of type $t$, and under channel condition $h$ being delivered successfully for all $n$. If $d_m(n)$ assumes the value 1 during slot $n$ when a packet from node $m$ is successfully *delivered*, and 0, otherwise, we have:

1. If $\boldsymbol{I}_m(n) = [1, 0, 0]$ then,

$$\begin{aligned} d_m(n) &= 0 \quad \ldots \text{with prob. } 1. \\ \implies \mathbb{E}[d_m(n)|\boldsymbol{I}_m(n) &= [1, 0, 0]] = 0. \end{aligned}$$

2. If $\boldsymbol{I}_m(n) = [0, 1, 0]$ then,

$$\begin{aligned} d_m(n) &= 1 \quad \ldots \text{with prob. } \mu_m = p^r_{1,m} p_{\text{chnl}} + p^r_{0,m}(1 - p_{\text{chnl}}) \\ \implies \mathbb{E}[d_m(n)|\boldsymbol{I}_m(n) &= [0, 1, 0]] = \mu_m. \end{aligned}$$

3. If $\boldsymbol{I}_m(n) = [0, 0, 1]$ then,

$$\begin{aligned} d_m(n) &= 1 \quad \ldots \text{with prob. } \nu_m = p^p_{1,m} p_{\text{chnl}} + p^p_{0,m}(1 - p_{\text{chnl}}) \\ \implies \mathbb{E}[d_m(n)|\boldsymbol{I}_m(n) &= [0, 1, 0]] = \nu_m. \end{aligned}$$

Combining all the above expectations using the indicator functions,

$$\mathbb{E}[d_m(n)|\boldsymbol{I}_m(n)] = \mu_m I_{\text{SR},m}(n) + \nu_m I_{\text{SP},m}(n). \tag{4.1}$$

Using the law of iterated expectations,

$$\begin{aligned} \mathbb{E}[d_m(n)] &= \mathbb{E}[\mathbb{E}[d_m(n)|\boldsymbol{I}_m(n)]] && (4.2) \\ &= \mathbb{E}[\mu_m I_{\text{SR},m}(n) + \nu_m I_{\text{SP},m}(n)] && (4.3) \\ &= \mu_m \mathbb{E}[I_{\text{SR},m}(n)] + \nu_m \mathbb{E}[I_{\text{SP},m}(n)] && (4.4) \\ &\triangleq \psi_m(n) && (4.5) \end{aligned}$$

**Note:-**

By construction of $d_m(n)$, we have, $\mathbb{E}[d_m(n)] = 1 p_{d_m(n)}(1) + 0 p_{d_m(n)}(0) = p_{d_m(n)}(1)$. Also, $d_m(n)$ indicates a successful transmission. Thus, $\psi_m(n)$ denotes the success probability of transmission of the $m^{\text{th}}$ node during slot $n$.

[1] We index the states to distinguish the interpretation of the states in different processes.

Further, the long-term throughput $\hat{q}_m$ is defined as:

$$\hat{q}_m = \mathbb{E}\left\{\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} d_m(n)\right\} \tag{4.6}$$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[d_m(n)] \tag{4.7}$$

$$= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\psi_m(n). \tag{4.8}$$

### 4.1.3 Possible Decisions by the Controller

For each node, the controller can instruct it to perform one of the following actions:(i) No sample - NS, (ii) Sample, transmit raw data - SR, or (iii) Sample, Process, transmit processed data - SP. Thus, the set $\mathcal{A} = \{\text{NS}, \text{SR}, \text{SP}\}$ is the action set.

**Definition 4.1.1: Indicator Variable Function**

The variable $I_{A,m}(n)$ is used to indicate when node $m$ is instructed by the controller to perform action $A$. Thus,

$$I_{A,m}(n) = \begin{cases} 1 & \text{if node takes action } A \text{ in slot } n, \\ 0 & \text{otherwise.} \end{cases}$$

The above definition is valid for all $A \in \mathcal{A}, m \in \mathcal{M}$.

The *action vector* for node $m$ for slot $n$ is given by $I_m(n) = [I_{\text{NS},m}(n), I_{\text{SR},m}(n), I_{\text{SP},m}(n)]$. With these definitions, we can develop the constraints to ensure that the decisions are consistent with the assumptions of the model. Consider,

$$I_{A,m}(n)\cdot(I_{A,m}(n)-1) = 0, \qquad \forall\, A \in \mathcal{A}, m \in \mathcal{M}, n \in \mathbb{N}, \tag{4.9a}$$

$$\mathbf{1}^{\top}\boldsymbol{I}_m(n) = 1, \qquad \forall\, m \in \mathcal{M}, n \in \mathbb{N}, \tag{4.9b}$$

$$\sum_{m=1}^{M} I_{\text{SR},m}(n) + I_{\text{SP},m}(n) \leqslant 1, \qquad \forall\, n \in \mathbb{N}. \tag{4.9c}$$

In the above, Equation 4.9a imposes the definition of the indicator variable, Equation 4.9b ensures that at most one action is performed by a node, and Equation 4.9c enforces a single transmission per slot.

#### Transmission Costs

Each of the actions in $\mathcal{A}$ is associated with an operating cost. This corresponding cost vector is given by $\boldsymbol{c} = [c_{\text{NS},m}, c_{\text{SR},m}, c_{\text{SP},m}]$ for node $m$. The real-time transmission cost function for node $m$ is given by:

$$C_m(n) = c_{\text{NS},m}I_{\text{NS},m}(n) + c_{\text{SR},m}I_{\text{SR},m}(n) + c_{\text{SP},m}I_{\text{SP},m}(n) \tag{4.10}$$

Thus, the constraint on the long-term time-averaged expected transmission cost is given by:

$$\bar{C}_m = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} C_m(n) \leqslant c_m, \forall\, m \in \mathcal{M}. \tag{4.11}$$

### 4.1.4 Cost of Actuation Error

**Definition 4.1.2: CAE Measure**

For the $m^{\text{th}}$ node, if $X_m$ and $\hat{X}_m$ denote the true state and the estimate during slot $n$, the Cost of Actuation Error is given by the function:

$$\Delta_m(n) = \text{CAE}(X_m = i, \hat{X}_m = j) = \delta_{i,j}^{(m)},$$

where, $i, j \in \mathcal{X}_m$. Note that in $\delta_{i,j}^{(m)}$, the term $(m)$ indicates the association with $m^{\text{th}}$ node and not the exponent.

In this work, we penalize out-of-sync ($i \neq j$) scenarios. Thus, we upper bound *average*[2] CAE when $\delta_{i,j} \geqslant 0$ if $i \neq j$ and $\delta_{i,j} \leqslant 0$ if otherwise. Using this *context-aware* distortion measure, we impose an upper bound on the long-term time-averaged expected CAE:

$$\bar{\Delta}_m = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Delta_m(n) \leqslant d_m, \forall\, m \in \mathcal{M}. \tag{4.12}$$

### 4.1.5 Age of Information

This work uses a general state-dependent form of the age that evolves as:

$$A_m(n+1) = \begin{cases} w_0 & \text{if } \hat{X}_m(n) = 0, d_m(n) = 1, \\ w_1 & \text{if } \hat{X}_m(n) = 0, d_m(n) = 1, \\ A_m(n) + 1 & \text{otherwise.} \end{cases} \tag{4.13}$$

To measure the freshness, we consider the weighted average AoI. Mathematically,

$$\frac{1}{M} \sum_{m=1}^{M} \alpha_m \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} A_m(n) = \lim_{N\to\infty} \frac{1}{MN} \sum_{n=1}^{N} \sum_{m=1}^{M} \alpha_m A_m(n), \tag{4.14}$$

where $\alpha_m \geqslant 0$ is the weight of the $m^{\text{th}}$ node.

## 4.2 Optimization Analysis

### 4.2.1 Problem Formulation

Putting the pieces together, we now present the main optimization problem this work address. Let $\mathcal{P}$ denote the class of all admissible policies.

[2]Unless mentioned otherwise, average here refers to the long-term time-averaged expected quantity.

**Problem 4: Main Problem**

$$\min_{\mathsf{P}\in\mathcal{P}} \quad \lim_{N\to\infty} \frac{1}{MN} \sum_{n=1}^{N} \sum_{m=1}^{M} \alpha_m \mathbb{E}[A_m(n)] \tag{4.15a}$$

$$\text{subject to} \quad \bar{\Delta}_m \leqslant d_m, \qquad \forall\, m \in \mathcal{M}, \tag{4.15b}$$

$$\bar{C}_m \leqslant c_m, \qquad \forall\, m \in \mathcal{M}, \tag{4.15c}$$

$$I_{A,m}(n)(I_{A,m}(n) - 1) = 0, \qquad \forall\, A \in \mathcal{A}, m \in \mathcal{M}, n \in \mathbb{N}, \tag{4.15d}$$

$$\mathbf{1}^\top \boldsymbol{I}_m(n) = 1, \qquad \forall\, m \in \mathcal{M}, n \in \mathbb{N}, \tag{4.15e}$$

$$\sum_{m=1}^{M} I_{\text{SR,m}}(n) + I_{\text{SR,m}}(n) \leqslant 1, \qquad \forall\, n \in \mathbb{N}. \tag{4.15f}$$

We will refer to Equation 4.15 as *main problem*, here on. In words, we wish to minimize the average weighted AoI (Equation 4.15a) subject to constraints on the CAE (Equation 4.15b), transmission costs (Equation 4.15c). The constraints Equation 4.15d-4.15f ensure the valid actions are prescribed by the policy.

### 4.2.2 Real-time expected CAE expression

Starting this part, we will borrow results from our basic work. The key observation that facilitates this direct borrowing is the definition of the variable $d_m(n)$ and the *action validity constraints*.

**Lemma 4.1**

The expression for real-time expected CAE is given by

$$\mathbb{E}[\Delta_m(n)] = \zeta_m + \xi_m \psi_m(n),$$

where $\zeta_m = \sum_{i,j} \delta_{i,j}^{(m)} \pi_i^{X_m} \pi_j^{X_m}$ and $\pi_0^{X_m} \pi_1^{X_m} \sum_{i,j} (-1)^{i+j} \delta_{i,j}^{X_m}$.

*Proof.* The proof is identical to the one done for single sensor set-up lemma 1.1. 😄

To be continued...

# Chapter 5

# Appendix

## 5.1 Distributions of the State and Estimate

**Theorem 5.1** Marginal and Joint Distribution of the true state and estimate

Let $P_{\hat{X}_n}$ and $P_{X_n,\hat{X}_n}$ denote the marginal and joint distribution of the estimate and estimate and true state respectively. Then $P_{\hat{X}_n} \sim \text{Bernoulli}(\pi_1^X)$ and

$$\begin{aligned} P(X_n = 0, \hat{X}_n = 0) &= (\pi_0^X)^2 + \pi_0^X \pi_1^X \psi(n), \\ P(X_n = 0, \hat{X}_n = 1) &= \pi_0^X \pi_1^X (1 - \psi(n)), \\ P(X_n = 1, \hat{X}_n = 0) &= pi_0^X \pi_1^X (1 - \psi(n)), \\ P(X_n = 1, \hat{X}_n = 1) &= (\pi_1^X)^2 + \pi_0^X \pi_1^X \psi(n). \end{aligned}$$

*Proof.* Assume that the system has attained steady-state. Let $S(n)$ represent the event of a successful update in a given time slot, which occurs whenever $d(n) = 1$. Let $\bar{S}(n)$ denote the complement of $S(n)$. We consider the marginal distribution of the estimate, specifically, $P(\hat{X}_n = 0)$ and $P(\hat{X}_n = 1)$. Using the law of total probability, we obtain

$$P(\hat{X}_n = 0) = P(\hat{X}_n = 0 \mid S(n))P(S(n)) + P(\hat{X}_n = 0 \mid \bar{S}(n))P(\bar{S}(n)). \tag{5.1}$$

Since a successful update resulting in $\hat{X}_n = 0$ can only occur if the true state was $X_n = 0$, we have

$$P(\hat{X}_n = 0 \mid S(n)) = P(X_n = 0) = \pi_0^X.$$

Furthermore, if there was no update (or an update attempt failed), the estimate $\hat{X}_n$ can be zero only if the previous estimate $\hat{X}_{n-1}$ was also zero, leading to

$$P(\hat{X}_n = 0 \mid \bar{S}(n)) = P(\hat{X}_{n-1} = 0).$$

The probabilities of a successful and unsuccessful update are given by $P(S(n)) = \psi(n)$ and $P(\bar{S}(n)) = 1 - \psi(n)$, respectively. Substituting these into equation 5.1, we obtain the recursive equation

$$P(\hat{X}_n = 0) = \pi_0^X \cdot \psi(n) + P(\hat{X}_{n-1} = 0) \cdot (1 - \psi(n)). \tag{5.2}$$

Now, let us examine a few iterations of this equation. Setting the initial (prior) distribution of $\hat{X}$ equal to that of $X$[1], i.e., letting $\hat{X}_0 \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(\pi_1^X)$.

For $n = 1$:

$$\begin{aligned} P(\hat{X}_1 = 0) &= \pi_0^X \cdot \psi(1) + P(\hat{X}_0 = 0) \cdot (1 - \psi(1)) &(5.3)\\ &= \pi_0^X \cdot \psi(1) + \pi_0^X \cdot (1 - \psi(1)) &(5.4)\\ &= \pi_0^X. &(5.5) \end{aligned}$$

[1] It was observed in during the simulations that the prior distribution had no impact on the steady-state behavior.

For $n = 2$:

$$P(\hat{X}_2 = 0) = \pi_0^X \cdot \psi(2) + P(\hat{X}_1 = 0) \cdot (1 - \psi(2)) \tag{5.6}$$
$$= \pi_0^X \cdot \psi(2) + \pi_0^X \cdot (1 - \psi(2)) \tag{5.7}$$
$$= \pi_0^X . \tag{5.8}$$

By induction, we conclude that $P(\hat{X}_n = 0) = \pi_0^X$ for all $n$. Parallely we also conclude $P(\hat{X}_n = 1) = \pi_1^X$. We now analyze the joint distribution $P_{X_n,\hat{X}_n}(i,j)$ for $i, j \in \{0,1\}$. Using the law of total probability,

$$P_{X_n,\hat{X}_n}(0,0) = P(X_n = 0, \hat{X}_n = 0) \tag{5.9}$$
$$= P(X_n = 0, \hat{X}_n = 0 \mid S(n))P(S(n)) + P(X_n = 0, \hat{X}_n = 0 \mid \bar{S}(n))P(\bar{S}(n)) \tag{5.10}$$
$$= P(\hat{X}_n = 0 \mid X_n = 0, S(n))P(X_n = 0 \mid S(n))P(S(n))$$
$$+ P(\hat{X}_n = 0 \mid X_n = 0, \bar{S}(n))P(X_n = 0 \mid \bar{S}(n))P(\bar{S}(n)). \tag{5.11}$$

Since an update ensures $\hat{X}_n$ aligns with $X_n$, we have $P(\hat{X}_n = 0 \mid X_n = 0, S(n)) = 1$. Further, if no update occurs, $\hat{X}_n = 0$ if and only if $\hat{X}_{n-1} = 0$, implying $P(\hat{X}_n = 0 \mid X_n = 0, \bar{S}(n)) = P(\hat{X}_{n-1} = 0)$. Substituting in equation 5.11,

$$P_{X_n,\hat{X}_n}(0,0) = \pi_0^X \cdot \psi(n) + P(\hat{X}_{n-1} = 0) \cdot \pi_0^X \cdot (1 - \psi(n)) \tag{5.12}$$
$$= \pi_0^X \cdot \psi(n) + \pi_0^X \cdot \pi_0^X \cdot (1 - \psi(n)) \tag{5.13}$$
$$= (\pi_0^X)^2 + \pi_0^X \cdot \pi_1^X \cdot \psi(n). \tag{5.14}$$

By symmetry, replacing 0 with 1 gives

$$P_{X_n,\hat{X}_n}(1,1) = (\pi_1^X)^2 + \pi_0^X \pi_1^X \psi(n). \tag{5.15}$$

Now, for the asymmetric case,

$$P_{X_n,\hat{X}_n}(0,1) = P(X_n = 0, \hat{X}_n = 1) \tag{5.16}$$
$$= P(X_n = 0, \hat{X}_n = 1 \mid S(n))P(S(n)) + P(X_n = 0, \hat{X}_n = 1 \mid \bar{S}(n))P(\bar{S}(n)) \tag{5.17}$$
$$= P(\hat{X}_n = 1 \mid X_n = 0, S(n))P(X_n = 0 \mid S(n))P(S(n))$$
$$+ P(\hat{X}_n = 1 \mid X_n = 0, \bar{S}(n))P(X_n = 0 \mid \bar{S}(n))P(\bar{S}(n)). \tag{5.18}$$

Since the estimate must match the true state in a time slot if we are given a successful update, $P(\hat{X}_n = 1 \mid X_n = 0, S(n)) = 0$, and by logic used previously, $P(\hat{X}_n = 1 \mid X_n = 0, \bar{S}(n)) = P(\hat{X}_{n-1} = 1)$, substituting back in equation 5.18,

$$P_{X_n,\hat{X}_n}(0,1) = 0 \cdot \pi_0^X \cdot \psi(n) + P(\hat{X}_{n-1} = 1) \cdot \pi_0^X \cdot (1 - \psi(n)) \tag{5.19}$$
$$= \pi_1^X \cdot \pi_0^X \cdot (1 - \psi(n)). \tag{5.20}$$

Similarly, interchanging 0 and 1 gives,

$$P_{X_n,\hat{X}_n}(1,0) = \pi_0^X \cdot \pi_1^X \cdot (1 - \psi(n)). \tag{5.21}$$

😄

## 5.2 Lower Bound on AoI-metric

To get a bound on the general expression of AoI-based metric, we will follow an approach that parallels that in [Kadota et al., 2018], with some modification.

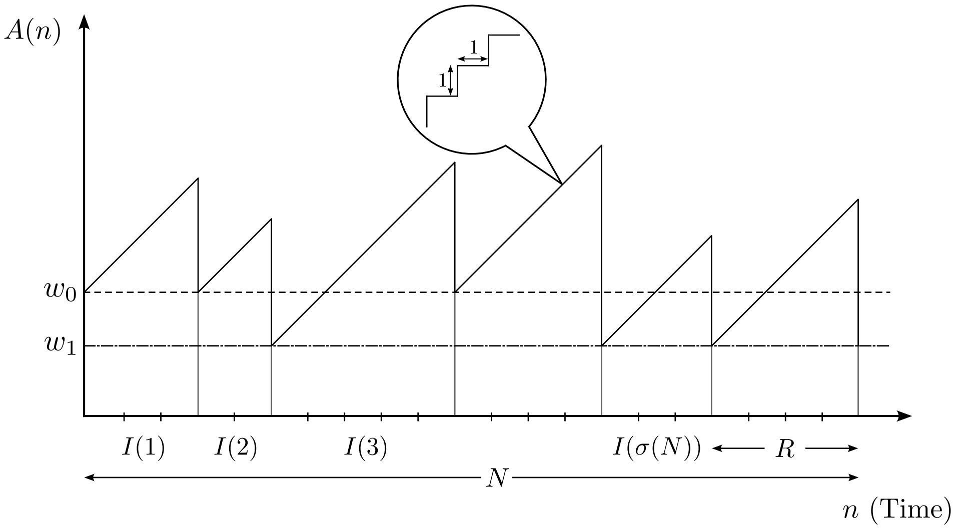



Figure 5.1: A representative sample path illustrating the time evolution of the Age of Information (AoI) over $N$ time slots is shown. The variable $I(k)$ denotes the inter-arrival time between the $(k-1)^{\text{th}}$ and $k^{\text{th}}$ packet arrivals. The quantity $\sigma(N)$ represents the total number of packets received within the observation window of length $N$.

**Theorem 5.2** Lower Bound on AoI-metric

The *long-term time-average of expected age* is lower bound by

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[A(n)] \geqslant \frac{1}{2}\left(\frac{1}{\hat{q}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0) \tag{5.22}$$

where, $\hat{q} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[d(n)]$ represents the long-term throughput of the source.

*Proof.* Consider a time-horizon of $n = 1, 2, 3, \ldots, N$. We define the following variables,

| **Variable** | **Description** |
|---|---|
| $\sigma(N)$ | Total number of samples successfully received at the end of $N$ slots. |
| $I(m)$ | Inter-arrival time of the $m^{\text{th}}$ sample/ packet. |
| $R$ | Number of left-over slots after last successful transmission. |

We also define the state indicator function of the $m^{\text{th}}$ sample,

$$s(m) \triangleq \begin{cases} 1 & \text{if the } m^{\text{th}} \text{ sample corresponds to state } 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.23}$$

**Note:-**

Observe that,

$$N = \sum_{m=1}^{\sigma(N)} \{I(m)\} + R. \tag{5.24}$$

We reiterate that the monitored process has reached a steady state, where it can be in either state in each time slot, with probabilities determined by the stationary distribution. By the definition of $s(m)$, we have,

$$s(m) = \begin{cases} 1 & \text{with probability } \pi_1^X \\ 0 & \text{with probability } 1 - \pi_1^X = \pi_0^X. \end{cases} \tag{5.25}$$

$$\therefore \mathbb{E}[s(m)] = 1 \cdot \pi_1^X + 0 \cdot \pi_0^X \tag{5.26}$$

$$= \pi_1^X \ \forall\, m. \tag{5.27}$$

For the $m^{\text{th}}$ packet, the age evolves as follows:

$$w_i, w_i + 1, w_i + 2, \ldots, w_i + I(m) - 1.$$

Here, $w_i$ depends on the state estimated from the previous sample. The total accumulated age for this packet, denoted as $A_m$, is given by

$$A_m = w_i + (w_i + 1) + (w_i + 2) + \cdots + (w_i + I(m) - 1) \tag{5.28}$$

$$= [w_i \cdot I(m)] + [1 + 2 + 3 + \cdots + (I(m) - 1)] \tag{5.29}$$

$$= (w_0 + s(m-1) \cdot (w_1 - w_0)) \cdot I(m) + \sum_{k=1}^{I(m)-1} k. \tag{5.30}$$

Using the formula for the sum of the first $(I(m) - 1)$ natural numbers, we obtain

$$A_m = \frac{(I(m) - 1)I(m)}{2} + (w_0 + s(m-1)(w_1 - w_0))\, I(m) \tag{5.31}$$

$$= \frac{I^2(m) - I(m)}{2} + (w_0 + s(m-1)(w_1 - w_0))\, I(m). \tag{5.32}$$

Summing the evolution of age across all received packets and accounting for the residual age after the last successfully received sample until slot $N$ for a single sample path (see figure 5.1), we obtain:

$$S = \sum_{n=1}^{N} A(n) = \sum_{m=1}^{\sigma(N)} A_m + A_R \tag{5.33}$$

$$= \sum_{m=1}^{\sigma(N)} \left[ \frac{I^2(m) - I(m)}{2} + (w_0 + s(m-1) \cdot (w_1 - w_0)) \cdot I(m) \right] + \frac{R^2 - R}{2} + (w_0 + s[\sigma(N)] \cdot (w_1 - w_0)) \cdot R \tag{5.34}$$

$$= \frac{1}{2} \left[ R^2 - R + \sum_{m=1}^{\sigma(N)} I^2(m) - \sum_{m=1}^{\sigma(N)} I(m) \right] + w_0 \cdot N + (w_1 - w_0) \cdot \left( s[\sigma(N)] \cdot R + \sum_{m=1}^{\sigma(N)} s(m-1) \cdot I(m) \right), \tag{5.35}$$

where $R$ denotes the left-over number of slots. Rearranging and combining the linear terms using equation 5.24 yields:

$$S = \frac{1}{2} \left[ R^2 + \sum_{n=1}^{\sigma(N)} I^2(m) - N \right] + w_0 \cdot N + (w_1 - w_0) \cdot \left( s[\sigma(n)] \cdot R + \sum_{m=1}^{\sigma(N)} s(m-1) \cdot I(m) \right). \tag{5.36}$$

**Definition 5.2.1: Sample Mean Operator**

We define the sample mean operator, $\mathbb{M}\{\cdot\}$ as :

$$\mathbb{M}\{I(m)\} = \frac{1}{\sigma(N)} \sum_{m=1}^{\sigma(N)} I(m)$$

$$\mathbb{M}\{I^2(m)\} = \frac{1}{\sigma(N)} \sum_{m=1}^{\sigma(N)} I^2(m)$$

$$S = \frac{1}{2}\left[R^2 + \sigma(N)\cdot\frac{1}{\sigma(N)}\sum_{m=1}^{\sigma(N)} I^2(m) - N\right] + w_0 N + (w_1 - w_0)\left(s[\sigma(N)]R + \sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right). \tag{5.37}$$

Applying Jensen's inequality to the sample mean operator, we obtain the following bound:

$$S \geqslant \frac{1}{2}\left[R^2 + \sigma(N)\left(\frac{1}{\sigma(N)}\sum_{m=1}^{\sigma(N)} I(m)\right)^2 - N\right] + w_0 N + (w_1 - w_0)\left(s[\sigma(N)]R + \sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right) \tag{5.38}$$

$$= \frac{1}{2}\left[R^2 + \frac{(N-R)^2}{\sigma(N)} - N\right] + (w_1 - w_0)\left(s[\sigma(N)]R + \sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right). \tag{5.39}$$

Taking the time average, we obtain:

$$\frac{1}{N}\sum_{n=1}^{N} A(n) \geqslant \frac{1}{2}\left[\frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)} - 1\right] + w_0 + \frac{(w_1 - w_0)}{N}\left(s[\sigma(N)]R + \sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right). \tag{5.40}$$

Now, defining $A$ as the time-average expected age over $N$ slots, we get:

$$A = \frac{1}{N}\mathbb{E}[S] \tag{5.41}$$

$$= \mathbb{E}\left[\frac{1}{2}\left(\frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)} - 1\right) + w_0 + \frac{(w_1 - w_0)}{N}\left(s[\sigma(N)]R + \sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right)\right]. \tag{5.42}$$

The expectation is taken with respect to the randomness in the communication channel, the true state of the process, the number of arrivals, and their inter-arrival times. By linearity of expectation,

$$A \geqslant \frac{1}{2}\left[\frac{R^2}{N} + \mathbb{E}\left[\frac{(N-R)^2}{N\sigma(N)}\right] - 1\right] + w_0 + \left(\frac{w_1 - w_0}{N}\right)\cdot\mathbb{E}\left[s[\sigma(n)]\cdot R + \sum_{m=1}^{\sigma(N)} s(m-1)\cdot I(m)\right] \tag{5.43}$$

$$= \frac{1}{2}\left[\frac{R^2}{N} + \mathbb{E}\left[\frac{(N-R)^2}{N\sigma(N)}\right] - 1\right] + w_0 + \left(\frac{w_1 - w_0}{N}\right)\cdot\left(\mathbb{E}\left[s[\sigma(n)]R\right] + \mathbb{E}\left[\sum_{m=1}^{\sigma(N)} s(m-1)I(m)\right]\right) \tag{5.44}$$

We now focus on the two expectation terms on the right-hand side. Consider the first term:

$$
\begin{aligned}
E_1 &= \mathbb{E}[s[\sigma(N)] \cdot R] && (5.45)\\
&= \mathbb{E}[s(\sigma(N))] \cdot \mathbb{E}[R] && \text{. . . Independence of } R \text{ and } s[\sigma(N)] \quad (5.46)\\
&= \pi_1^X \cdot \mathbb{E}[R] && \text{. . . From Equation (5.27).} \quad (5.47)
\end{aligned}
$$

Next, we analyze the second expectation term:

$$
E_2 = \mathbb{E}\left[\sum_{m=1}^{\sigma(N)} s(m-1) \cdot I(m)\right]. \tag{5.48}
$$

Since the expectation is taken with respect to the randomness in $\sigma(\cdot)$, $s(\cdot)$, and $I(\cdot)$, we apply the law of iterated expectation:

$$
\begin{aligned}
E_2 &= \mathbb{E}_{\sigma(N)}\left[\mathbb{E}\left[\sum_{m=1}^{\sigma(N)} s(m-1) \cdot I(m) \middle| \sigma(N)\right]\right] && (5.49)\\
&= \mathbb{E}_{\sigma(N)}\left[\sum_{m=1}^{\sigma(N)} \mathbb{E}[s(m-1) \cdot I(m) | \sigma(N)]\right] && \text{. . . Linearity of Expectation} \quad (5.50)\\
&= \mathbb{E}_{\sigma(N)}\left[\sum_{m=1}^{\sigma(N)} \mathbb{E}[s(m-1)] \cdot \mathbb{E}[I(m)] \middle| \sigma(N)\right] && \text{. . . Independence of } s(m-1) \text{ and } I(m) \quad (5.51)\\
&= \mathbb{E}_{\sigma(N)}\left[\sum_{m=1}^{\sigma(N)} \pi_1^X \cdot \mathbb{E}[I(m)] \middle| \sigma(N)\right] && \text{. . . From equation 5.27} \quad (5.52)\\
&= \pi_1^X \cdot \mathbb{E}_{\sigma(N)}\left[\sum_{m=1}^{\sigma(N)} \mathbb{E}[I(m)] \middle| \sigma(N)\right] && \text{. . . Homogeneity of Expectation} \quad (5.53)\\
&= \pi_1^X \cdot \mathbb{E}_{\sigma(N)}\left[\mathbb{E}\left(\sum_{m=1}^{\sigma(N)} I(m) \middle| \sigma(N)\right)\right] && \text{. . . Linearity of Expectation} \quad (5.54)\\
&= \pi_1^X \cdot \mathbb{E}\left[\sum_{m=1}^{\sigma(N)} I(m)\right] && \text{. . . Law of Iterated Expectation} \quad (5.55)\\
&= \pi_1^X \cdot \mathbb{E}\left[N - R\right] && \text{. . . From } N = \sum_{m=1}^{\sigma(N)} I(m) + R \quad (5.56)\\
&= \pi_1^X \cdot (N - \mathbb{E}[R]) && \text{. . . Linearity of Expectation.} \quad (5.57)
\end{aligned}
$$

Going back to our original inequality and re-substituting $E_1$ and $E_2$,

$$
\begin{aligned}
A &\geqslant \frac{1}{2}\mathbb{E}\left[\frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)} - 1\right] + w_0 + \left(\frac{w_1 - w_0}{N}\right) \cdot \left(\pi_1^X \cdot \mathbb{E}[R] + \pi_1^X \cdot (N - \mathbb{E}[R])\right) && (5.58)\\
&= \frac{1}{2}\mathbb{E}\left[\frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)} - 1\right] + w_0 + \left(\frac{w_1 - w_0}{N}\right) \cdot (\pi_1^X \cdot N) && (5.59)\\
&= \frac{1}{2}\mathbb{E}\left[\frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)} - 1\right] + w_0 + \pi_1^X \cdot (w_1 - w_0). && (5.60)
\end{aligned}
$$

To analytically minimize the argument of the expectation, define

$$
f(R) = \frac{R^2}{N} + \frac{(N-R)^2}{N\sigma(N)}.
$$

Noting that $R \geqslant 0$ and $f$ is monotonically increasing in $R$,

$$f(R) \geqslant \frac{(N^2)}{N\sigma(N)} \geqslant \frac{N}{1+\sigma(N)}. \tag{5.61}$$

Applying this inequality along with the linearity of expectation:

$$\frac{1}{N}\sum_{n=1}^{N} A(n) \geqslant \frac{1}{2}\mathbb{E}\left[\frac{N}{1+\sigma(N)} - 1\right] + w_0 + \pi_1^X \cdot (w_1 - w_0) \tag{5.62}$$

$$= \frac{1}{2}\left(N\mathbb{E}\left[\frac{1}{1+\sigma(N)}\right] - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0). \tag{5.63}$$

Applying Jensen's inequality to $g(x) = \frac{1}{x}, x \in \mathbb{R}_{++}$, followed by linearity of expectation:

$$\frac{1}{N}\sum_{n=1}^{N} A(n) \geqslant \frac{1}{2}\left(\frac{1}{\frac{1}{N} + \frac{1}{N}\mathbb{E}[\sigma(N)]} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0). \tag{5.64}$$

Taking the limit as $N \to \infty$,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} A(n) \geqslant \frac{1}{2}\left(\frac{1}{\lim_{N\to\infty} \frac{1}{N} + \frac{1}{N}\mathbb{E}[\sigma(N)]} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0) \tag{5.65}$$

$$= \frac{1}{2}\left(\frac{1}{\lim_{N\to\infty} \frac{1}{N}\mathbb{E}[\sigma(N)]} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0). \tag{5.66}$$

Since:

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}[\sigma(N)] = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{n=1}^{N} d(n)\right] \tag{5.67}$$

$$= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[d(n)]. \tag{5.68}$$

By the definition of long-term throughput equation 1.17, we obtain:

$$\bar{A} = \lim_{N\to\infty} A \geqslant \frac{1}{2}\left(\frac{1}{\hat{q}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0). \tag{5.69}$$

☺

## 5.3 Long-term time-average of expected AoI under SRP

**Theorem 5.3** Long-term time-average of expected AoI under SRP

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[A(n)]^{\mathsf{R}} = \left(\frac{1}{\hat{q}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0).$$

*Proof.* We can interpret the possible AoI evolutions as sample-paths of a discrete-time Markov chain (DTMC) in itself. Without loss of generality, we assume that $w_0 < w_1$. The state space of this DTMC is given by

$$\Omega_A = \{w_0, w_0 + 1, \ldots, w_1 - 1, w_1, w_1 + 1, \ldots\}. \tag{5.70}$$

Since, under the stationary randomized policy (SRP), decisions are made independently of any factor, we expect the transition probabilities to be independent of the current state. We proceed by deriving the stationary distribution of this Markov chain, followed by computing its expected value.

### 5.3.1 Transition Probabilities

Consider any $s \in \Omega_A$,

$$P(a_{n+1} = w_0 \mid a_n = s) = P(X_n = 0, S(n)) \tag{5.71}$$

$$= P(X_n = 0)P(S(n)) \quad \text{...(Independence of events).} \tag{5.72}$$

We know that $P(X_n = 0) = \pi_0^X$[2] and $P(S(n)) = \psi(n)^{\mathsf{R}} = \psi^{\mathsf{R}}$. Thus,

$$P(a_{n+1} = w_0 \mid a_n = s) = \pi_0^X \cdot \psi^{\mathsf{R}}. \tag{5.73}$$

Similarly, we can show that

$$P(a_{n+1} = w_1 \mid a_n = s) = \pi_1^X \cdot \psi^{\mathsf{R}}, \tag{5.74}$$

except when $s = w_1 - 1$. Using these, let us define:

$$p_0 = P(a_{n+1} = w_0 \mid a_n = s), \tag{5.75}$$

$$p_1 = P(a_{n+1} = w_1 \mid a_n = s), \tag{5.76}$$

$$p_2 = P(a_{n+1} = s + 1 \mid a_n = s). \tag{5.77}$$

Since these probabilities must sum to one, we have:

$$p_0 + p_1 + p_2 = 1, \tag{5.78}$$

$$\therefore p_2 = 1 - p_0 - p_1, \tag{5.79}$$

$$= 1 - \pi_0^X \cdot \psi^{\mathsf{R}} - \pi_1^X \cdot \psi^{\mathsf{R}}, \tag{5.80}$$

$$= 1 - \left(\pi_0^X + \pi_1^X\right) \cdot \psi^{\mathsf{R}}, \tag{5.81}$$

$$= 1 - \psi^{\mathsf{R}}. \tag{5.82}$$

| **Quantity** | **Variable** |
|---|---|
| Type, Channel-dependent success probabilities. | $p_0^r, p_0^t, p_1^r,$ and $p_1^p$ |
| Probability of sampling of a type of sample. | $p_{\text{SR}}, p_{\text{SP}}$ |
| Probability of channel being good. | $p_{\text{chnl}}$ |
| Stationary Distribution of the external process. | $\pi_0^X, \pi_1^X$ |
| Stationary Distribution of the AoI metric. | $\pi_i^A$ |
| Transition Matrix of the external process. | $\boldsymbol{P}_X$ |
| Transition Matrix of the AoI metric. | $\boldsymbol{P}_A$ |

Table 5.1: Reference Table for some of the variables and their meanings.

Using these transition probabilities,we can now draw a state transition diagram. Refer figure 5.2. Its corresponding state transition matrix is given by, $\boldsymbol{P}_A$,

$$\boldsymbol{P}_A = \begin{bmatrix} \pi_0^X \cdot \psi^{\mathsf{R}} & 1-\psi^{\mathsf{R}} & 0 & \cdots & 0 & \pi_1^X \cdot \psi^{\mathsf{R}} & 0 & \cdots \\ \pi_0^X \cdot \psi^{\mathsf{R}} & 0 & 1-\psi^{\mathsf{R}} & \cdots & 0 & \pi_1^X \cdot \psi^{\mathsf{R}} & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \\ \pi_0^X \cdot \psi^{\mathsf{R}} & 0 & 0 & \cdots & 1-\psi^{\mathsf{R}} & \pi_1^X \cdot \psi^{\mathsf{R}} & 0 & \cdots \\ \pi_0^X \cdot \psi^{\mathsf{R}} & 0 & 0 & \cdots & 0 & 1-\psi^{\mathsf{R}}+\pi_1^X \cdot \psi^{\mathsf{R}} & 0 & \cdots \\ \pi_0^X \cdot \psi^{\mathsf{R}} & 0 & 0 & \cdots & 0 & \pi_1^X \cdot \psi^{\mathsf{R}} & 1-\psi^{\mathsf{R}} & \cdots \\ \pi_0^X \cdot \psi^{\mathsf{R}} & 0 & 0 & \cdots & 0 & \pi_1^X \cdot \psi^{\mathsf{R}} & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \end{bmatrix} \tag{5.83}$$

[2]This quantity represents the stationary distribution of the two-state Markov chain used to model the external phenomenon.

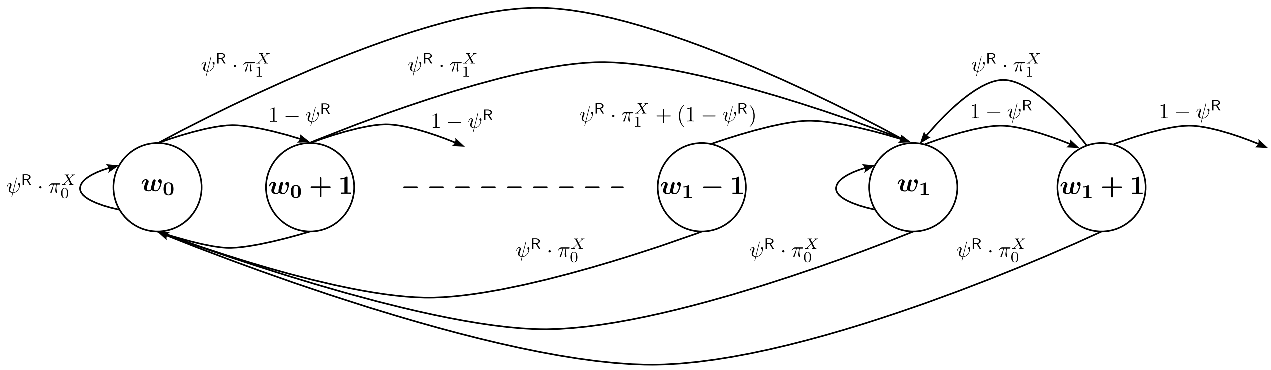



Figure 5.2: State Transition Diagram of the AoI metric Markov chain.

$$
\begin{array}{c|cccccccc}
 & w_0 & w_0+1 & & & w_1-1 & w_1 & w_1+1 & \\
w_0 & \pi_0^{(X)}\cdot\psi & 1-\psi & 0 & \cdots & 0 & \pi_1^{(X)}\cdot\psi & 0 & \cdots \\
w_0+1 & \pi_0^{(X)}\cdot\psi & 0 & 1-\psi & \cdots & 0 & \pi_1^{(X)}\cdot\psi & 0 & \cdots \\
 & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \\
w_1-2 & \pi_0^{(X)}\cdot\psi & 0 & 0 & \cdots & 1-\psi & \pi_1^{(X)}\cdot\psi & 0 & \cdots \\
w_1-1 & \pi_0^{(X)}\cdot\psi & 0 & 0 & \cdots & 0 & 1-\psi+\pi_1^{(X)}\cdot\psi & 0 & \cdots \\
w_1 & \pi_0^{(X)}\cdot\psi & 0 & 0 & \cdots & 0 & \pi_1^{(X)}\cdot\psi & 1-\psi & \cdots \\
w_1+1 & \pi_0^{(X)}\cdot\psi & 0 & 0 & \cdots & 0 & \pi_1^{(X)}\cdot\psi & 0 & \cdots \\
 & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots
\end{array}
$$

Figure 5.3: Construction of the transition matrix of age metric.

The figure 5.3 explains the construction of this matrix. Using normalization condition to ensure that the laws of probability are satisfied,

$$\sum_{s=w_0}^{\infty} \pi_s^A = 1. \tag{5.84}$$

We solve for the stationary distribution, $\boldsymbol{\pi}_A$, using the above and the equation $\boldsymbol{\pi}_A = \boldsymbol{\pi}_A \boldsymbol{P}_A$. This results in the following equations:

$$
\begin{aligned}
\pi^A_{w_0} &= \pi^A_{w_0} \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} + \pi^A_{w_0+1} \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} + \dots \\
&= \big[ \sum_{s=w_0}^{\infty} \pi^A_s \big] \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} \\
&= \pi^X_0 \cdot \psi^{\mathsf{R}} && (5.85) \\
\pi^A_{w_0+1} &= \pi^A_{w_0} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}}) && (5.86) \\
\pi^A_{w_0+2} &= \pi^A_{w_0+1} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^2 && (5.87) \\
&\vdots \\
\pi^A_{w_1-1} &= \pi^A_{w_1-2} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{w_1-1-w_0} && (5.88) \\
\pi^A_{w_1} &= \pi^A_{w_0} \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} + \pi^A_{w_0+1} \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} + \dots + \pi^A_{w_1-1} \cdot [1-\psi^{\mathsf{R}} + \pi^X_1 \cdot \psi^{\mathsf{R}}] + \pi^A_{w_1} \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} + \dots \\
&= \big[ \sum_{s=w_0}^{\infty} \pi^A_s \big] \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} + \pi^A_{w_1-1} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_1 \cdot \psi^{\mathsf{R}} + \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{w_1-w_0} && (5.89) \\
\pi^A_{w_1+1} &= \pi^A_{w_1} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_1 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}}) + \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{w_1+1-w_0} && (5.90) \\
\pi^A_{w_1+2} &= \pi^A_{w_1+1} \cdot (1-\psi^{\mathsf{R}}) \\
&= \pi^X_1 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^2 + \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{w_1+2-w_0} && (5.91) \\
&\vdots
\end{aligned}
$$

Compactly expressing the result,

$$
\pi^A_k = \begin{cases} \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{k-w_0} & \text{if } w_0 \leqslant k \leqslant w_1 - 1, \\ \pi^X_0 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{k-w_0} + \pi^X_1 \cdot \psi^{\mathsf{R}} \cdot (1-\psi^{\mathsf{R}})^{k-w_1} & \text{if } w_1 \leqslant k. \end{cases} \tag{5.92}
$$

**Expectation of Age**

$$
\begin{aligned}
\mathbb{E}[A(n)] &= \sum_{k=w_0}^{\infty} k \cdot P(A(n)=k) && (5.93) \\
&= \sum_{k=w_0}^{\infty} k \cdot \pi^A_k && (5.94) \\
&= \sum_{k=w_0}^{w_1-1} k \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_0} + \sum_{k=w_1}^{\infty} k \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_0} + k \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_1} && (5.95) \\
&= \sum_{k=w_0}^{w_1-1} k \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_0} + \sum_{k=w_1}^{\infty} k \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_0} + \sum_{k=w_1}^{\infty} k \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_1} && (5.96) \\
&= \sum_{k=w_0}^{\infty} k \cdot \pi^X_0 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_0} + \sum_{k=w_1}^{\infty} k \cdot \pi^X_1 \cdot \psi^{\mathsf{R}} (1-\psi^{\mathsf{R}})^{k-w_1} && (5.97)
\end{aligned}
$$

**Claim 5.3.1** $\sum_{k=s}^{\infty} k \cdot \theta^{k-s} = \frac{\theta}{(1-\theta)^2} + \frac{s}{1-\theta}$

*Proof.* Consider,

$$S = \sum_{k=s}^{\infty} k \cdot \theta^{k-s} \tag{5.98}$$

Making the substitution $k - s \to k$,

$$\begin{aligned} S &= \sum_{k=0}^{\infty} (k+s) \cdot \theta^{k} && (5.99) \\ &= \sum_{k=0}^{\infty} k \cdot \theta^{k} + s \cdot \sum_{k=0}^{\infty} \theta^{k} && (5.100) \\ &= \theta \cdot \sum_{k=0}^{\infty} k \cdot \theta^{k-1} + \frac{s}{1-\theta} && (5.101) \\ &= \theta \cdot \sum_{k=0}^{\infty} \frac{d\theta^{k}}{d\theta} + \frac{s}{1-\theta} && (5.102) \\ &= \theta \cdot \frac{d}{d\theta} \Big( \sum_{k=0}^{\infty} \theta^{k} \Big) + \frac{s}{1-\theta} && (5.103) \\ &= \theta \cdot \frac{d}{d\theta} \Big( \frac{1}{1-\theta} \Big) + \frac{s}{1-\theta} && (5.104) \\ \therefore S &= \frac{\theta}{(1-\theta)^2} + \frac{s}{1-\theta} && (5.105) \end{aligned}$$

😄

Using the above result,

$$\begin{aligned} \mathbb{E}[A(n)] &= \pi_0^X \cdot \psi^{\mathsf{R}} \cdot \Big[\frac{1-\psi}{(\psi^{\mathsf{R}})^2} + \frac{w_0}{\psi^{\mathsf{R}}}\Big] + \pi_1^X \cdot \psi^{\mathsf{R}} \cdot \Big[\frac{1-\psi^{\mathsf{R}}}{(\psi^{\mathsf{R}})^2} + \frac{w_1}{\psi^{\mathsf{R}}}\Big] && (5.106) \\ &= \pi_0^X \cdot \frac{1-\psi^{\mathsf{R}} + w_0 \cdot \psi^{\mathsf{R}}}{\psi^{\mathsf{R}}} + \pi_1^X \cdot \frac{1-\psi^{\mathsf{R}} + w_1 \cdot \psi^{\mathsf{R}}}{\psi^{\mathsf{R}}} && (5.107) \\ &= \frac{w_0 \cdot (\pi_0^X \cdot \psi^{\mathsf{R}}) + w_1 \cdot (\pi_1^X \cdot \psi^{\mathsf{R}}) + (1-\psi^{\mathsf{R}})}{\psi^{\mathsf{R}}} && (5.108) \\ &= \left(\frac{1}{\psi^{\mathsf{R}}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0) && (5.109) \end{aligned}$$

Thus, the long-term time-averaged expected AoI under SRP is given as,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[A(n)] = \left(\frac{1}{\psi^{\mathsf{R}}} - 1\right) + w_0 + \pi_1^X \cdot (w_1 - w_0) \tag{5.110}$$

😄

# Bibliography

[Fountoulakis et al., 2023] Fountoulakis, E., Pappas, N., and Kountouris, M. (2023). Goal-oriented policies for cost of actuation error minimization in wireless autonomous systems. *IEEE Communications Letters*, 27(9):2323–2327.

[Kadota et al., 2018] Kadota, I., Sinha, A., and Modiano, E. (2018). Optimizing Age of Information in Wireless Networks with Throughput Constraints. In *IEEE INFOCOM*, pages 1844–1852.

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

[S. et al., 2023] S., J., Pappas, N., and Bhat, R. V. (2023). Distortion Minimization with Age of Information and Cost Constraints. In *WiOpt*, pages 1–8, Singapore, Singapore.

[Yates et al., 2021] Yates, R. D., Sun, Y., Brown, D. R., Kaul, S. K., Modiano, E., and Ulukus, S. (2021). Age of Information: An Introduction and Survey. *IEEE J. Sel. Areas Commun.*, 39(5):1183–1210.