Fresh Data Delivery: Joint Sampling and Routing for Minimizing the Age of Information

Adem Utku Atasayar
atasayar.utku@metu.edu.tr
Middle East Technical University
Ankara, Turkiye

Aimin Li
liaimingogo@gmail.com
Middle East Technical University
Ankara, Turkiye

Çağrı Arı
ari.cagri@metu.edu.tr
Middle East Technical University
Ankara, Turkiye

Elif Uysal
uelif@metu.edu.tr
Middle East Technical University
Ankara, Turkiye

Abstract

In this paper, we extend the *freshness*-oriented sampling problem by incorporating controlled delay statistics through *heterogeneous routing options*, where the Age of Information (AoI) serves as the metric for data *freshness*. Our objective is to jointly optimize sampling and routing policies to minimize the *long-term average* AoI, where the sender can choose to forward each status update over one of the available routes, which have distinct delay statistics. This problem is an infinite-horizon Semi-Markov Decision Process (SMDP) with an *uncountable* state space and a *hybrid* action space, consisting of discrete routing choices and continuous waiting times. We develop an efficient algorithm to solve this problem and theoretically establish that the optimal policy exhibits a threshold structure, characterized by: (i) a threshold-based *monotonic handover mechanism* for optimal routing, where the switching order aligns with the decreasing order of mean delays; and (ii) a multi-threshold *piecewise linear waiting mechanism* for optimal sampling, where the total number of thresholds is upper bounded by $2N-1$, given N selectable routes. We implement the proposed algorithm in a *satellite-terrestrial* integrated routing scenario, and simulation results reveal an intriguing insight: **routes with higher average delay or variance can still contribute to minimizing AoI**.

CCS Concepts

• **Information systems**; • **Networks** → **Network design and planning algorithms**; *Network performance analysis*;

Keywords

Age of Information, Markov Decision Process, Sampling policy, Routing policy, Threshold policy

ACM Reference Format:
Adem Utku Atasayar, Aimin Li, Çağrı Arı, and Elif Uysal. 2025. Fresh Data Delivery: Joint Sampling and Routing for Minimizing the Age of Information. In *International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc '25), October 27–30, 2025, Houston, TX, USA*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3704413.3764413

ACM ISBN 979-8-4007-1353-8/25/10
https://doi.org/10.1145/3704413.3764413

1 Introduction

1.1 Background and Related Work

In an increasingly connected world where systems rely on remotely sampled data to make real-time decisions, the freshness of data samples has become a key driver of application performance. Hence, information freshness is emerging as a Key Performance Indicator (KPI) in next-generation communication networks spanning wired, wireless, and non-terrestrial links. For instance, in remote-sensing-based emergency response systems, access to *fresh* data regarding environmental variables supports real-time risk assessment and enhances response efficiency. Similarly, in Vehicle to Everything (V2X) scenarios, vehicles rely on continuously updated information collected through multi-sensor data fusion to navigate the rapidly changing environmental conditions. Moreover, in Industrial Internet of Things (IIoT) applications, the staleness of sensor data negatively impacts production efficiency and operational safety.

This growing emphasis on information freshness has led to the development of the Age of Information (AoI) metric that quantifies it [11]. Distinctly from packet delay, AoI serves as a flow-based performance indicator, providing a receiver-centric view of information staleness. Formally, at any given time t, the AoI at the destination is defined as $\Delta(t) = t - U(t)$, where $U(t)$ denotes the generation timestamp of the most recently received data sample. Clearly, a low age requires a sufficiently high throughput of data updates that observe sufficiently low latency. Hence, AoI combines the conventional performance metrics of latency and throughput in a novel and interesting way.

However, AoI is a simple notion in the sense that it focuses only on the freshness of data. While simple, its minimization has been subject to a rich set of problems related to minimizing age and monotone functions thereof [22, 24] under energy constraints [1, 2, 26], and more complex network settings, including multi-hop [23], multi-source [3–5, 9], and broadcast networks [10]. Furthermore, in scenarios where the significance of data extends beyond temporal freshness, more sophisticated evaluation frameworks have been developed [6, 7, 13, 18, 19, 21]. Some of these frameworks utilize AoI as an intermediate metric to capture task-specific relevance through the freshness of data samples [18, 19].

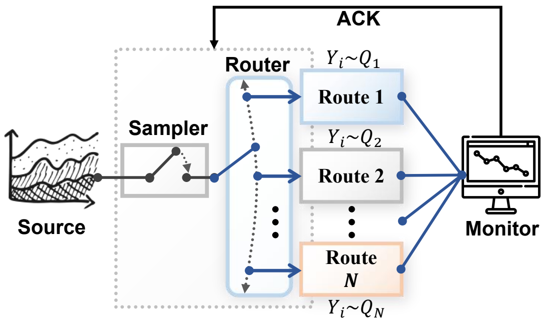

Figure 1: A remote monitoring system, where status updates are transmitted through N heterogeneous routes.

In this paper, we focus on the optimization of AoI. Our goal is to extend the formulation of Age-optimal sampling and scheduling first proposed in [22] to a case where there are multiple routing options between the source and the destination. The new formulation proposed in this paper was inspired and motivated by the growing interest in Satellite IoT and integrated TN-NTN in 5G and 6G, where data transmission decisions are sometimes faced with a choice between routing through non-terrestrial links versus terrestrial connections. Routing through non-terrestrial connections typically entails long RTTs due to high propagation delay and delay variability caused by intermittent availability of inter-satellite links, and sometimes resorting to a store-and-forward mode. On the other hand, terrestrial connections may be prone to congestion and outages. Our formulation addresses the fundamental question of how route selection impacts information freshness in these hybrid terrestrial/non-terrestrial environments, providing a theoretical foundation for age-aware joint sampling and routing in next-generation communication networks.

1.2 Contributions of This Work

• **System Model**: We formulate a novel joint sampling and routing problem in which the transmitter optimizes both the *sampling time* and the *data forwarding route* to minimize the *long-term average* AoI at the destination. Our formulation is a direct extension of the problem in [22], and in contrast to this and other prior work where the sampling problem is attacked under given delay statistics [7, 13, 21, 24], our system model takes a *proactive* approach by actively selecting and switching routes to control the delay experienced by packets. Meanwhile, unlike existing multi-channel scheduling problems that either focus on homogeneous channels with uniform one-slot delays [20] or heterogeneous channels where each channel experiences an on-off constant discrete delay [14, 15], our work considers distinct *continuous delay distributions* across different routes and jointly optimizes sampling and route selection strategies.

• **Solution Methodology**: We show that this sampling and routing co-design problem is a Semi-Markov Decision Process (SMDP) with an *uncountable* state space and a *hybrid* action space, consisting of discrete routing choices and continuous waiting times. Such SMDPs are known to be challenging to solve due to the size of the state and action spaces, and previous research has addressed this complexity by: (i) discretizing the *uncountable* state space and the action space [12, 28], which introduces quantization error or; (ii) focusing on a special case of SMDP where the state transitions are independent of actions[1] [21, 22, 24], which, however, does not hold in our scenario. In this paper, we develop a new algorithm namely ***R****elative* ***E****xpected* ***A****ction* ***V****alue* ***I****teration* (REAVI) that directly solves this class of SMDPs without discretizing the space. *To the best of our knowledge, this is the first algorithm that handles such SMDPs while avoiding quantization errors.*

• **Structural Results**: We prove that each of the jointly optimal sampling and routing policies exhibit a graceful *threshold structure*: (i) The routing policy is a thresholds-based handover policy, where a specific route is selected when the current AoI at the receiver falls within certain range, precisely determined by multiple thresholds; (ii) A new sample is taken and transmitted when the AoI at the receiver reaches a threshold that is a function of the selected transmission route. These structural properties deem the policies suitable for practical implementation. We designed an efficient algorithm to determine these thresholds.

• **Counter-Intuitive Insights:** We test our algorithms on the model of an *integrated satellite-terrestrial* communication network scenario. Our simulation results reveal an intriguing insight: routes with higher mean delay and/or greater variance can still contribute to minimizing AoI. This finding challenges conventional wisdom that may prioritize routes with minimal mean delay characteristics. It demonstrates that the strategic utilization of diverse routing options in complex network environments can lead to superior information freshness.

2 Model and Formulation

2.1 System Model

We consider a remote monitoring system, as illustrated in Fig. 1, consisting of a source, a sampler, a router, and a monitor. Status updates are timely generated, and each is transmitted through one of the N heterogeneous routes, with the objective of maintaining the freshest possible information on the monitor at all times. We assume that each route $k \in \mathcal{N} \triangleq \{1, \cdots, N\}$ has a stationary transmission delay distribution, denoted by Q_k, with finite first and second moments. Specifically, the mean delay is given by $\mu_k \triangleq \mathbb{E}_{Q_k}[Y] < \infty$, and the variance by $\sigma_k^2 \triangleq \mathbb{E}_{Q_k}[(Y-\mu_k)^2] < \infty$, for all $k \in \mathcal{N}$. We also assume a non-preemptive system, where a new transmission can begin only after the previous one has been completed. Upon receiving each data sample, the monitor sends an ideal acknowledgment (ACK) to the transmitter, indicating that the system is ready to initiate the next transmission.

We adopt the *generate-at-will* model [22, 26], in which the sampler can become active at any time, provided that a new transmission is allowed. We next introduce some notation. After receiving the ACK corresponding to the previous transmission, the $(i+1)$-th data sample is generated and submitted to route R_i at time instant S_{i+1}. It is subsequently delivered to the monitor at time instant $D_{i+1} = S_{i+1} + Y_{i+1}$, where $Y_{i+1} \sim Q_{R_i}$ denotes the random transmission delay experienced by the $(i+1)$-th sample. The initial conditions of the system are set as follows: A sample is submitted

[1]This allows the Markov decision process to be reduced to a renewal reward process, thus simplifying the analysis.

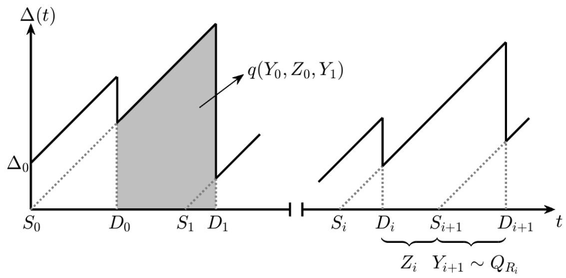

Figure 2: Sample evolution of the AoI process $\Delta(t)$.

to route k, which is arbitrarily selected from the N available routes, at time instant $S_0 = 0$. Consequently, the corresponding delivery time instant is $D_0 = Y_0$, where $Y_0 \sim Q_k$.

2.2 Age of Information

The Age of Information (AoI) is the metric of our interest to measure the *freshness* of information. This metric is defined as the time elapsed since the generation of the most recently received data sample [11]. Specifically, the AoI $\Delta(t)$ at time t is defined by[2]

$$\Delta(t) \triangleq t - S_i, \quad \text{if } D_i \leq t < D_{i+1}. \tag{1}$$

The AoI $\Delta(t)$ is a stochastic process that increases linearly over time and experiences downward jumps to the most recent delay value Y_i upon the delivery of the i-th data sample at time D_i, as illustrated in Fig. 2. The value of $\Delta(t)$ between the time instants $S_0 = 0$ and D_0 is assumed to increase linearly, starting from an arbitrary finite initial real value $\Delta(0) = \Delta_0 < \infty$.

2.3 Problem Formulation

We aim at minimizing the long-term average AoI by designing a joint sampling and routing policy $\pi \triangleq (R_0, Z_0, R_1, Z_1 \dots)$. This policy consists of two distinct sequences: (i) a sequence of routing decisions $(R_0, R_1, R_2, \dots)$, where R_i specifies the route selected for transmitting the $(i+1)$-th packet, and (ii) a sequence of finite waiting times $(Z_0, Z_1, Z_2, \dots)$, where $Z_i < \infty$ determines the $(i+1)$-th sampling (or submission) time as $S_{i+1} = D_i + Z_i$. Let Π denote the set of all causal and stationary policies π. The corresponding optimization problem is then formulated as follows:

Problem 1 (Average Age Minimization).

$$\lambda^\star \triangleq \min_{\pi \in \Pi} \limsup_{T \to \infty} \mathbb{E}\left[\frac{1}{T}\int_0^T \Delta(t)\, dt\right], \tag{2}$$

where $\lambda^\star$ denotes the optimal long-term average AoI.

2.4 Dinkelbach's Reformulation

The AoI process $\Delta(t)$ is inherently a piecewise linear function as defined by equation (1). Hence, it is natural to rewrite Problem 1 as an average sum over time intervals $[D_i, D_{i+1})$

$$\limsup_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\int_{t=0}^T \Delta(t)dt\right] = \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1})\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]}, \tag{3}$$

where $q(Y_i, Z_i, Y_{i+1})$ is the integral of $\Delta(t)$ from D_i to D_{i+1}:

$$\begin{aligned} q(Y_i, Z_i, Y_{i+1}) &= \int_{D_i}^{D_{i+1}} (t - S_i)dt = \int_{Y_i}^{Y_i+Z_i+Y_{i+1}} \tau d\tau \\ &= \frac{1}{2}(Y_i + Z_i + Y_{i+1})^2 - \frac{1}{2}Y_i^2. \end{aligned} \tag{4}$$

The region corresponding to the integral $q(Y_0, Z_0, Y_1)$, defined over the time interval $[D_0, D_1)$, is illustrated in Fig. 2.

Using equation (3), Problem (1) can be reformulated as follows:

$$\lambda^\star \triangleq \min_{\pi\in\Pi} \limsup_{n\to\infty} \frac{\mathbb{E}\left[\sum_{i=0}^{n-1} q(Y_i, Z_i, Y_{i+1})\right]}{\mathbb{E}\left[\sum_{i=0}^{n-1}(Z_i + Y_{i+1})\right]}. \tag{5}$$

The problem in equation (5) is an infinite-horizon Semi-Markov Decision Process (SMDP) with an *uncountable* state space and a hybrid action space, making it particularly challenging to solve. To address this, we propose a solution approach in two steps. In the first step, we apply Dinkelbach's method for nonlinear fractional programming to eliminate the fractional structure of the objective function in equation (5). Consequently, for any given $\lambda \geq 0$, we consider the following problem.

Problem 2 (Dinkelbach's Reformulation).

$$h(\lambda) \triangleq \min_{\pi\in\Pi} \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}\left[q(Y_i, Z_i, Y_{i+1})\right] - \lambda\mathbb{E}\left[(Z_i + Y_{i+1})\right]. \tag{6}$$

The following Lemma 1 establishes the relationship between the function $h(\lambda)$ and the optimal long-term average AoI $\lambda^\star$.

Lemma 1. *The following assertions hold true:*
(1) $\lambda^\star \gtreqless \lambda$ *if and only if* $h(\lambda) \gtreqless 0$.
(2) *If* $h(\lambda) = 0$*, the solutions to* (5) *and* (6) *are identical.*

Proof. See Appendix A. □

According to Lemma 1, the solution to the problem in equation (5) can be obtained by identifying the value of λ for which $h(\lambda) = 0$ and solving Problem 2 for that specific λ. The root of the function $h(\lambda)$ is exactly the optimal long-term average age $\lambda^\star$.

2.5 Average Cost MDP for a Given λ

In the second step, we show that Problem 2 can be cast as an average-cost Markov Decision Process (MDP) for a given value of λ, described by the quadruple $\mathscr{M}(\lambda) \triangleq (\mathcal{S}, \mathcal{A}, \mathcal{P}, C)$:

- State Space $\mathcal{S} = [0, \infty)$: The state of the MDP is $Y_i = y \in \mathcal{S}$.
- Action Space $\mathcal{A} = \mathbb{R}^+ \times \mathcal{N}$: The action of the MDP is represented by a tuple $(R_i, Z_i) \in \mathcal{A}$, where $R_i = r \in \mathcal{N}$ is the routing decision and $Z_i = z < \infty$ is the waiting time.
- State Transition Probability $\mathcal{P} : \mathcal{S} \times \mathcal{A} \to [0, 1]$: The state transition probability $P(Y_{i+1} \in \mathcal{S}' \mid Y_i = y, R_i = r, Z_i = z)$ under the action $(R_i, Z_i) = (r, z) \in \mathcal{A}$, for any measurable subset $\mathcal{S}' \subseteq \mathcal{S}$, is given by

$$\int_{\mathcal{S}'} f(y' \mid y, r, z)\, dy' = \int_{\mathcal{S}'} Q_r(y')\, dy', \tag{7}$$

[2]The framework in this work can be readily extended to general AoI penalty functions $g(\Delta(t))$ as alternative metrics of interest, where $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ represents any monotonically non-decreasing function. However, due to page limits, we focus on the linear case in this paper.

where $f(y' \mid y, r, z)$ denotes the probability density function (PDF) of the next state Y_{i+1} given the current state $Y_i = y$ and action $(R_i, Z_i) = (r, z)$, and $Q_r(y')$ is the PDF of the transmission delay associated with route r.

- Cost Function $C : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$: The real-valued cost function defined over the state and action spaces, denoted by $g(Y_i, Z_i, R_i; \lambda)$, can be expressed as follows by inspecting the objective function of Problem 2:

$$g(y, z, r; \lambda) = \mathbb{E}_{Q_r}\left[q(y, z, Y_{i+1})\right] - \lambda(z + \mathbb{E}_{Q_r}[Y_{i+1}])$$
$$= \mathbb{E}_{Q_r}\left[\frac{(2y + Y_{i+1} + z)(Y_{i+1} + z)}{2}\right] - \lambda z - \lambda \mathbb{E}_{Q_r}[Y_{i+1}]$$
$$= \frac{z^2}{2} + (y + \mu_r - \lambda)z + (y - \lambda)\mu_r + \frac{\mu_r^2 + \sigma_r^2}{2}. \tag{8}$$

Using Lemma 1 and the constructed MDP $\mathscr{M}(\lambda)$ for a given λ, we can develop a *nested* two-layer algorithm, e.g., [3, 13], to solve the SMDP in Problem (1). The details of this numerical solution are presented in Section 4.

3 Main Results

Before presenting the numerical solution to Problem 1, we first establish several structural results regarding the jointly optimal sampling and routing policies.

3.1 Structural Results of Optimal Policies

The following Theorem 1 establishes the piecewise-threshold structure of the jointly optimal sampling and routing policies.

Theorem 1. *For an N-route problem where the mean delay of each route satisfies $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_N$ and the delay distribution of each route has infinite support, the jointly optimal sampling and routing policies exhibit the following threshold structure:*

(1) *Optimal Routing: The optimal routing action at the i-th epoch, denoted by $R_i^\star$, is a monotonic non-decreasing step function of the observed delay Y_i, and can be determined by $K \leq N - 1$ positive thresholds $0 < \tau_1 < \tau_2 \cdots < \tau_K$ and $K + 1$ monotonic increasing index values $a_1 < a_2 < \cdots < a_{K+1} \in \mathcal{N}$:*

$$R_i^\star = \sum_{k=1}^{K+1}(a_k - a_{k-1})u(Y_i - \tau_{k-1}), \tag{9}$$

where $\tau_0 \triangleq 0$, $a_0 \triangleq 0$, and $u(t)$ is the unit step function:

$$u(t) \triangleq \begin{cases} 0, & t < 0 \\ 1, & t \geq 0. \end{cases} \tag{10}$$

(2) *Optimal Sampling: The optimal waiting time at the i-th epoch, denoted by $Z_i^\star$, follows a water-filling structure and can be determined by $K + 1$ thresholds $\beta_1^\star < \beta_2^\star < \cdots < \beta_{K+1}^\star$ with $\beta_k^\star = \lambda^\star - \mu_{a_k}$,*

$$Z_i^\star = \begin{cases} (\beta_1^\star - Y_i)^+, & 0 \leq Y_i < \tau_1 \\ \vdots & \vdots \\ (\beta_K^\star - Y_i)^+, & \tau_{K-1} \leq Y_i < \tau_K \\ (\beta_{K+1}^\star - Y_i)^+, & \tau_K \leq Y_i \end{cases}, \tag{11}$$

or equivalently,

$$Z_i^\star = (\lambda^\star - \mu_{R_i^\star} - Y_i)^+. \tag{12}$$

where $\lambda^\star$ is the optimal average AoI defined in Problem 1, and $(\cdot)^+$ is defined as $(\cdot)^+ \triangleq \max\{0, \cdot\}$.

Proof Sketch. With the MDP $\mathscr{M}(\lambda)$ (8), we can establish the *Average-Cost Optimality Equation* (ACOE) [8, Eq. 4.1]:

$$V^*(y; \lambda) + h(\lambda) = \min_{z,r}\left\{g(y, z, r; \lambda) + \mathbb{E}_{Q_r}\left[V^*(Y_{i+1}; \lambda)\right]\right\}, \tag{13}$$

where $V^*(y; \lambda)$ is the relative value function, and $h(\lambda)$ is the optimal value of the reformulated MDP in Problem 2. Given any λ and route $r \in \mathcal{N}$, we first prove that the optimal waiting time that solves the right hand-side of (13) follows a *water-filling* structure, given by:

$$z^\star(y; r, \lambda) = (\lambda - \mu_r - y)^+. \tag{14}$$

As $h(\lambda^\star) = 0$, applying $\lambda = \lambda^\star$ in (13) and (14) yields:

$$V^*(y; \lambda^\star) = \min_r\left\{g\big(y, z^\star(y; r, \lambda^\star), r; \lambda^\star\big) + \mathbb{E}_{Q_r}\left[V^*(Y_{i+1}; \lambda^\star)\right]\right\}. \tag{15}$$

For short-hand notations, we define the action-value function as:

$$Q(y, r) \triangleq g\big(y, z^\star(y; r, \lambda^\star), r; \lambda^\star\big) + \mathbb{E}_{Q_r}\left[V^*(Y_{i+1}; \lambda^\star)\right], \tag{16}$$

and the optimal routing policy $r^\star(y)$ turns to

$$r^\star(y) = \arg\min_r\{Q(y, r)\}. \tag{17}$$

Then, we analyze a series of properties of the function $Q(y, r)$ and prove that $r^\star(y)$ is a non-decreasing step function, thus accomplishing the proof. See Section 5 for the detailed proof. □

We have established that the AoI-optimal routing policy follows a threshold-based structure. However, one can argue that these thresholds never actually exist and that the optimal policy always uses a single route. To counter this, we show that there exist system configurations where these thresholds must exist.

Lemma 2 (Existence of The Thresholds). *Consider a system with two routes whose delay distributions have infinite support, and suppose their means and standard deviations satisfy: $\mu_1 > \mu_2 > 0$ and $0 < \sigma_1 < \sigma_2$. Then, for any given distribution $Q_2(\mu_2, \sigma_2)$, there exists values (μ_1, σ_1) such that a threshold τ_1 exists.*

Proof. See Appendix D. □

It is also worth noting that even when $\lambda_1 \geq \lambda_2$[3] the age-optimal policy might still use route 1 for some range of y, depending on how the action-value functions $Q(y, r)$'s behave. Hence, τ_1 may still exist even in such cases.

The following Lemma 3 demonstrates an important relationship between the sampling threshold $\beta_i^\star$ and the routing threshold τ_i.

Lemma 3. *The following assertion holds true:*

$$\beta_k^\star < \tau_k, \quad k \in \{1, \ldots, K\}. \tag{18}$$

Proof. We know from Lemma 6 that the optimal route j at $y = \tau_i$ satisfies $\mu_j < \mu_{a_i}$. Then, we have

$$\frac{\partial Q(y, j)}{\partial y} = \frac{\partial Q(y, a_i)}{\partial y} = \lambda^\star - y, \quad y < \beta_i^\star. \tag{19}$$

Therefore, τ_i must be greater than $\beta_i^\star$. □

[3] λ_k denotes the average AoI achieved by only using route k with the AoI-optimal sampling policy provided in [22, Theorem 4], as defined in Appendix D.

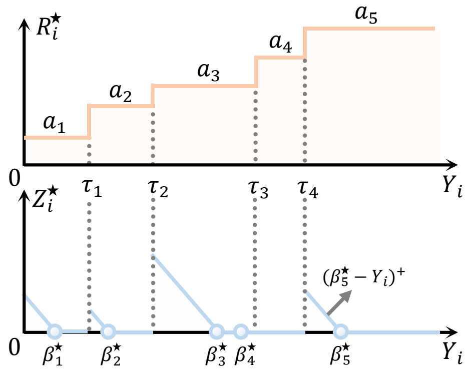

Figure 3: Visualization of the jointly optimal policies.

Consequently, for any interval $Y_i \in [\tau_{k-1}, \tau_k)$ associated with a constant optimal routing option a_k, there exists a corresponding sub-interval $[\beta_k^\star, \tau_k)$ in which a zero-waiting policy, defined by $Z_i^\star = (\beta_k^\star - Y_i)^+ = 0$, is optimal. An example sketch for the structure of the jointly optimal sampling and routing policy is depicted in Fig. 3. The threshold-based structure derived in this subsection enables highly efficient deployment in complex networks. Terminals can maximize information *freshness* simply by storing and applying the derived thresholds. In Section 4, we present a series of algorithms to compute these thresholds efficiently.

3.2 When is it Optimal to use a Single Route?

At first glance, it may seem natural to choose the route with the lowest average delay to ensure fresh information delivery. Nevertheless, Theorem 1 demonstrates that, in general, a combination of $K+1$ heterogeneous routes contributes to minimizing the AoI. In this subsection, we examine the sufficient and necessary condition in which a single-route policy is optimal.

Theorem 2 (Optimality of Single Route Policy). *Consider the N-route problem given in Theorem 1, the following assertions hold true:*
(1) If the route k policy[4] is optimal, it must have $\mu_k = \mu_N$.
(2) The route k policy is the optimal policy if and only if

$$\arg\min_r Q(0, r) = k, \tag{20}$$

where $Q(y, r)$ is defined in (16).

Proof. See Appendix E □

3.3 Bounds on the Optimal Average AoI

In this subsection, we establish the upper and lower bounds on the optimal average age $\lambda^\star$. These bounds will later serve as initialization points for the bisection search described in Section 4.3.

Lemma 4. *$\lambda^\star$ is upper and lower bounded by:*

$$0 \le \lambda^\star \le \min_r \left\{ \frac{3\mu_r}{2} + \frac{\sigma_r^2}{2\mu_r} \right\}. \tag{21}$$

Proof. See Appendix F. □

[4] We call a single route policy which always uses route k as route k policy.

4 Numerical Solutions

We propose algorithms to solve Problem (1) and determine the thresholds introduced in Theorem 1. By leveraging the Dinkelbach's method discussed in 2.3, this problem can be solved by a *nested* two-layer structure: In the inner layer, we approximate $h(\lambda)$ in the ACOE (13) for each λ; in the outer layer, we conduct bisection search to approach the optimal value $\lambda^\star$ where $h(\lambda^\star) = 0$.

4.1 Challenges in Approximating $h(\lambda)$

4.1.1 Challenge 1: Hybrid Action Space. The first challenge to solve $h(\lambda)$ arises from the *hybrid* action space, where routing decisions $r \in \mathcal{N}$ are discrete and waiting actions $z \in \mathbb{R}^+$ are *uncountable*. To address this, we leverage (14), which establishes the optimal sampling policy $z^\star(y, r; \lambda) = (\lambda - \mu_r - y)^+$. Substituting this into the ACOE (13) yields a transformed SMDP with a simplified *countable* routing action space:

$$V^*(y;\lambda) + h(\lambda) = \min_r \left\{ g(y, z^\star(y; r, \lambda), r; \lambda) + \mathbb{E}_{Q_r}\left[V^*(Y_{i+1}; \lambda) \right] \right\}. \tag{22}$$

This transformation effectively decouples the action space and reduces the dimension of the action space.

4.1.2 Challenge 2: Uncountable State Space. The second challenge in solving for $h(\lambda)$ stems from the *uncountable* nature of the state space. The ACOE (22) must hold for any state $y \in \mathbb{R}^+$, making it computationally infeasible to evaluate the value function at every point in the continuous state space $\mathcal{S}$. Conventional approaches address this challenge by discretizing the continuous state space into a finite grid with M regions $\{y_1, y_2, \cdots, y_M\}$, where each element y_i represents a distinct region of the state space $\mathcal{S}$ [12, 27, 28]. With this discretization, the Relative Value Iteration (RVI) algorithm [25] can be applied to approximate $h(\lambda)$.

While theoretically sound, this space grid approach introduces a quantization error ϵ_M. Although this error asymptotically vanishes as M increases, i.e., $\lim_{K\to\infty} \epsilon_K = 0$, this improvement comes at the cost of substantially increased computational complexity. As the discretization becomes finer (i.e., as M increases), the computational burden grows quadratically. This is because each iteration of the RVI algorithm has a computational complexity of $O(NM^2)$, where N is the number of routes available for transmission.

4.2 Proposed ReaVI Algorithm

To avoid this undesirable trade-off between quantization error and high computation complexity, we propose a new algorithm named **R**elative **E**xpected **A**ction **V**alue **I**teration (ReaVI) that eliminates the need for state space discretization. First, inspired by the typical RVI algorithm [25], we establish the relative value function:

$$W^*(y;\lambda) \triangleq V^*(y;\lambda) - V^*(0;\lambda), y \in \mathbb{R}^+, \tag{23}$$

where $y = 0$ serves as the reference state with $W^*(0, \lambda) = 0$. Substituting (23) into (22) yields a new optimality function:

$$W^*(y;\lambda) + h(\lambda) = \min_r \left\{ g(y, z^\star(y; r, \lambda), r; \lambda) + \mathbb{E}_{Q_r}\left[W^*(Y_{i+1}; \lambda) \right] \right\}, \tag{24}$$

where $h(\lambda)$ is obtained by setting $y = 0$:

$$h(\lambda) = \min_r \left\{ g(0, z^\star(0; r, \lambda), r; \lambda) + \mathbb{E}_{Q_r}\left[W^*(Y_{i+1}; \lambda) \right] \right\}. \tag{25}$$

Algorithm 1: *Relative Expected Action Value Iteration (ReaVI)*

Input: MDP tuple $\mathscr{M}(\lambda) = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{C})$, tolerance $\epsilon > 0$

Initialize:
$G(r;\lambda) \leftarrow 0, \quad \forall r \in \mathcal{N}$; // REAV Initialization
$h(\lambda) \leftarrow \min_r\{g(0, z^\star(r, 0), r; \lambda)\}$; // $h(\lambda)$ Initialization
repeat
  $h_{\text{old}}(\lambda) \leftarrow h(\lambda)$; // Store previous value
  for $q \in \mathcal{N}$ **do**
    $G(q;\lambda) \leftarrow -h_{\text{old}}(\lambda) + \mathbb{E}_{Y \sim Q_q}\left[\min_r \left\{g(Y, z^\star(r, Y), r; \lambda) + G(r, \lambda)\right\}\right]$;
    // REAV Function Update Based on REAVOE (27)
  $h(\lambda) \leftarrow \min_r\{g(0, z^\star(r, 0), r; \lambda) + G(r;\lambda)\}$; // Update $h(\lambda)$ Based on (27)
until $|h(\lambda) - h_{\text{old}}(\lambda)| < \epsilon$;
Output: $h(\lambda)$ and $\{G(r;\lambda)\}_{r\in\mathcal{N}}$

Algorithm 2: *Route Thresholds Algorithm*

Input: $\lambda^\star, \{G(r;\lambda^\star)\}_{r\in\mathcal{N}}$

Initialize:
$i \leftarrow 1$; // Iteration number initialization
$a_1 = \arg\min_r\{g(0, z^\star(r, 0), r; \lambda^\star) + G(r, \lambda^\star)\}$; // Optimal route for $y = 0$
$a_{\text{old}} \leftarrow a_1$; // Store initial value
repeat
  $\tau_i \leftarrow \infty$;
  for $r > a_{\text{old}}$ **do**
    $x_r \leftarrow \{y : Q(y, a_{old}) = Q(y, r)\}$; // Find the unique crossover point
    if $x_r \leq \tau_i$ **then**
      $\tau_i \leftarrow x_r$; // Update minimum threshold
      $a_{i+1} \leftarrow r$; // Update optimal route
  $a_{old} \leftarrow a_{i+1}$; // Update newly found optimal route
  $i = i + 1$; // Set for next iteration
until $a_i = N$;
Output: $\{\tau_j\}_{j\in\{1,\dots,i-1\}}$ and $\{a_j\}_{j\in\{1,\dots,i\}}$

However, this problem remains challenging due to the *uncountable* nature of $\mathbb{R}^+$. To address this challenge, we introduce a new function namely **R**elative **E**xpected **A**ction **V**alue (REAV) function, which is:

$$G(r;\lambda) \triangleq \mathbb{E}_{Q_r}\left[W^*(Y_{i+1};\lambda)\right], r \in \mathcal{N}. \tag{26}$$

This transformation shifts our focus from optimizing over the *uncountable* state space to optimizing over the finite set of route decisions. By randomizing (24) with $y = Y$, where $Y \sim Q_q$ and applying the expectation operator $\mathbb{E}_{Q_q}[\cdot]$ to both sides of (24), we derive a new average cost optimality equation for $\forall q \in \mathcal{N}$:
***R**elative **E**xpected **A**ction **V**alue **O**ptimality **E**quation* (REAVOE):

$$G(q;\lambda) + h(\lambda) = \mathbb{E}_{Y\sim Q_q}\left[\min_r \left\{g\big(Y, z^\star(Y; r, \lambda), r; \lambda\big) + G(r, \lambda)\right\}\right], \tag{27}$$

where $h(\lambda)$ is given by substituting (26) into (25):

$$h(\lambda) = \min_r \left\{g\big(0, z^\star(0; r, \lambda), r; \lambda\big) + G(r;\lambda)\right\}. \tag{28}$$

This developed equation converts an *infinite-dimensional* problem with *uncountable* state space $\mathbb{R}^+$ into a *finite-dimensional fixed-point equation* over the *countable* action space $\mathcal{N}$, and thus eliminates the quantization error. To solve the REAVOE, we propose an algorithm based on fixed-point iteration, as shown in Algorithm 1.

4.3 Bisec-ReaVI

After obtaining the approximated optimal value $h(\lambda)$ from the ReaVI algorithm, we implement a bisection search to identify the optimal average AoI $\lambda^\star$ where $h(\lambda^\star) = 0$. The theoretical foundation for this approach is established in Lemma 1, which proves a fundamental monotonic relationship: $\lambda^\star \gtreqless \lambda$ if and only if $h(\lambda) \gtreqless 0$. This strict monotonicity guarantees the convergence of the bisection method to the unique optimal long-term average AoI $\lambda^\star$.

4.4 Route Thresholds Solutions

Given the bisection-derived value $\lambda^\star$ and the corresponding REAV functions $\{G(r;\lambda^\star)\}_{r\in\mathcal{N}}$ obtained from Algorithm 1, we can establish the closed-form action-value function $Q(y, r) = G(r;\lambda^\star) + g(y, z^\star(y, r;\lambda^\star), r; \lambda^\star)$. A direct way to identify the thresholds in Theorem 1 is to discretize $\mathcal{S}$ into grid points $\{y_k\}_{k=1}^M$ and compute $\min_r Q(y_k, r)$ at each point. Each threshold τ_i is then defined as the breakpoint where the minimizing r value changes: $\arg\min_r Q(y_{m+1}, r) - \arg\min_r Q(y_m, r) \geq 1$. The complexity of such an algorithm is $O(MN)$, where M must be sufficiently large to capture all thresholds.

To overcome this inefficiency, we are motivated to establish the *Route Thresholds Algorithm* shown in Algorithm 2, which avoids scanning all M points and achieves $O(N^2)$ worst-case complexity, further dropping to $O(N)$ if the optimal solution involves only a single route.

5 Proof of Optimal Threshold Structures

For short-hand notations, we define $Q(y, z, r; \lambda)$ as the *state-action function* in the right-hand side of (15):

$$Q(y, z, r; \lambda) \triangleq g(y, z, r; \lambda) + \mathbb{E}_{Q_r}[V^*(Y_{i+1}; \lambda)]. \tag{29}$$

Given a specific route r for $Q(y, z, r; \lambda)$, we first solve the conditionally optimal $z^\star(y; r, \lambda)$.

- Case 1: If $\lambda - \mu_r - y \leq 0$, we have that
$$\frac{\partial Q(y, z, r; \lambda)}{\partial z} = z + y + \mu_r - \lambda \geq 0. \tag{30}$$
In this case, $Q(y, z, r; \lambda)$ is monotonically increasing with z given a specific r and y, which indicates that $z^\star(y; r, \lambda) = 0$.
- Case 2: If $\lambda - \mu_r - y > 0$, from (30) we can establish that if $z \in (0, \lambda - \mu_r - y)$, $Q(y, z, r; \lambda)$ is monotonically decreasing

with z; if $z \in [\lambda - \mu_r - y, \infty)$, $Q(y, z, r; \lambda)$ is monotonically increasing with z. As a result, $z^\star(y; r, \lambda) = \lambda - \mu_r - y$.

Combining the aforementioned two cases yields:

$$z^\star(y; r, \lambda) = (\lambda - \mu_r - y)^+. \tag{31}$$

Substituting (31) into (29) and setting $\lambda = \lambda^\star$ yields a compact form of $Q(y, r)$, whose definition has been given in (16):

$$\begin{aligned} Q(y, r) = &- \frac{\left((\lambda^\star - \mu_r - y)^+\right)^2}{2} + (y - \lambda^\star)\mu_r \\ &+ \frac{\sigma_r^2 + \mu_r^2}{2} + \mathbb{E}_{Q_r}\left[V^*(Y_{i+1}; \lambda^\star)\right]. \end{aligned} \tag{32}$$

With the notation $Q(y, r)$, the ACOE turns to:

$$V^*(y; \lambda^\star) = \min_r \{Q(y, r)\}, y \in \mathbb{R}^+. \tag{33}$$

Meanwhile, the optimal routing policy is given by:

$$r^\star(y) = \arg\min_r \{Q(y, r)\}. \tag{34}$$

To analyze the threshold structure of $r^\star(y)$, the following lemma discusses some important properties of $Q(y, r)$ and $V^*(y; \lambda^\star)$.

Lemma 5. *The following assertions hold true:*

(1) *$\forall r \in \mathcal{N}$, $Q(y, r)$ is monotonically increasing with y.*
(2) *$V^*(y; \lambda^\star)$ is monotonically increasing with y.*
(3) *For any routes j, k such that $\mu_j > \mu_k$, we have*

$$\frac{\partial Q(y, j)}{\partial y} \geq \frac{\partial Q(y, k)}{\partial y}, \forall y \in \mathbb{R}^+. \tag{35}$$

Proof. See Appendix B. □

With (1) and (3) of Lemma 5 in hand, we can then establish the following lemma, which indicates that the optimal routing policy $r^\star(y)$ is monotonically non-decreasing with y:

Lemma 6. *Consider N routes with their mean delays satisfying $\mu_1 \geq \mu_2 \cdots \geq \mu_N$, if route j is optimal at $y = y^\star$, we have that*

$$\begin{aligned} r^\star(y) &= \arg\min_r \{Q(y, r)\} \geq j, \text{if } y > y^\star, \\ r^\star(y) &= \arg\min_r \{Q(y, r)\} \leq j, \text{if } y < y^\star. \end{aligned} \tag{36}$$

Proof. Since route j is optimal at $y = y^\star$, we have

$$Q(y^\star, j) \leq Q(y^\star, i), \tag{37}$$

for any i. Now, for $i < j$ we know $\mu_i \geq \mu_j$. Then, combining (35) with (37) we obtain

$$Q(y, j) \leq Q(y, i), \ y \geq y^\star, \tag{38}$$

which proves that no route $i < j$ can be optimal for $y > y^\star$. The proof for the converse statement follows the same logic. □

As Lemma 6 holds for $\forall y, y^\star$, $r^\star(y)$ is a monotonically non-decreasing function with respect to y. As $r^\star(y)$ belongs to a discrete set $\mathcal{N}$, it forms a non-decreasing step function as shown in (9). Substituting the step function $r^\star(y)$ into (31) yields:

$$z^\star(y) = z^\star(y; r^\star(y), \lambda^\star) = (\lambda^\star - \mu_{r^\star(y)} - y)^+. \tag{39}$$

For a given constant-value interval $[\tau_{k-1}, \tau_k)$ where $r^\star(y) = a_k$, the optimal sampling policy is defined as:

$$z^\star(y) = (\lambda^\star - \mu_{a_k} - y)^+. \tag{40}$$

Defining $\beta_k^\star \triangleq \lambda^\star - \mu_{a_k}$, we next prove that the *water-filling levels* $\beta_k^\star$ are strictly increasing with the index k. First, we can show that for any $i < j$, it follows that $\mu_{a_i} \leq \mu_{a_j}$, which leads to

$$\beta_i^\star = (\lambda^\star - \mu_{a_i})^+ \leq (\lambda^\star - \mu_{a_j})^+ = \beta_j^\star. \tag{41}$$

Next, we prove that $\beta_i^\star \neq \beta_j^\star$ for $i \neq j$. This is achieved by the following lemma, which indicates that $\mu_{a_i} \neq \mu_{a_j}$ for $i \neq j$.

Lemma 7. *Let $\mathcal{R}^\star = \{a_1, \ldots, a_{K+1}\}$ denote the set of routes used by the age-optimal policy and let $\mathcal{G}_\mu$ be defined as*

$$\mathcal{G}_\mu \triangleq \{r \in \mathcal{N} : \mu_r = \mu\}. \tag{42}$$

Then, at most one route from $\mathcal{G}_\mu$ can belong to the optimal set $\mathcal{R}^\star$:

$$|\mathcal{R}^\star \cap \mathcal{G}_\mu| \leq 1, \forall \mu \in \mathbb{R}^+. \tag{43}$$

Proof. See Appendix C. □

With lemma 7 and (41), we establish that $\beta_1^\star < \cdots < \beta_{K+1}^\star$.

6 Simulation Results

This section presents simulation results for practical scenarios to validate the analytical findings and evaluate the performance of our proposed algorithm.

6.1 Comparing Benchmarks

In this subsection, we refer to our designed jointly optimal sampling and routing policy as the "optimal policy" and evaluate its performance against the following benchmark policies:

• *Minimum Average Delay Routing with AoI-Optimal Sampling (**MAD-Optimal**)*: This policy always selects the route with the minimum average delay. Given this selection, the AoI-optimal waiting strategy from [22, Theorem 4] is implemented to minimize the long-term average AoI.

• *Minimum Delay Variance Routing with Zero-Wait Sampling (**MDV-Zero Wait**)*: This policy consistently selects the route with the lowest delay variance. It is combined with a zero-wait strategy, where a new packet is sampled and transmitted immediately upon the delivery of the previous packet[5].

• *Minimum Delay Variance Routing with AoI-Optimal Sampling (**MDV-Optimal**)*: This policy always selects the route with the minimum delay variance and follows the AoI-optimal waiting strategy as outlined in [22, Theorem 4].

6.2 Satellite-Terrestrial Integrated Routes

We consider two distinct classes of routes, denoted by $\mathcal{N}_{\text{Sat}}$ and $\mathcal{N}_{\text{Ter}}$. Here, $\mathcal{N}_{\text{Sat}}$ represents the set of Low Earth Orbit (LEO) Satellite routes with stochastic delays, while $\mathcal{N}_{\text{Ter}}$ represents the set of terrestrial routes with stochastic delays.

6.2.1 LEO Satellite Routes with Stochastic Delays. For $l \in \mathcal{N}_{\text{Sat}}$, the delay is modeled by a *log-normal distribution*, characterized by the following probability density function [16]:

$$P_{Y \sim Q_l}(y) = \frac{1}{y\beta_l\sqrt{2\pi}} \exp\left(-\frac{(\ln y - \alpha_l)^2}{2\beta_l^2}\right), l \in \mathcal{N}_{\text{Sat}}, \tag{44}$$

[5]Zero-wait policy [22] is work-conserving, hence, it achieves maximum throughput on any given route. MAD-Zero Wait was also included in the simulations; however, its significantly suboptimal performance led us to omit it from the figures for simplicity.

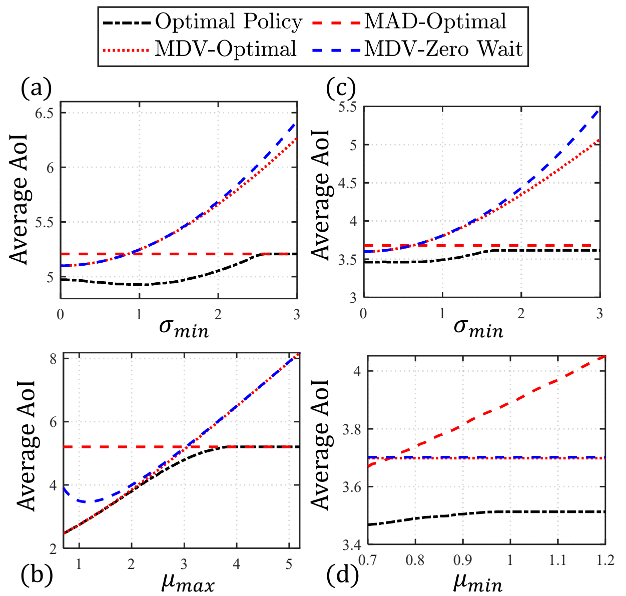

Figure 4: Simulation results of systems with $N = 2$ and $N = 3$.

Table 1: Simulation Parameters

Route	Route 1	Route 2	Route 3
Distribution Parameters	Log-normal (μ_1, σ_1)	Gamma (μ_2, σ_2)	Gamma (μ_3, σ_3)
Fig. 4 (a)	$(3.4, [0, 3])$	$(0.7, 5)$	—
Fig. 4 (b)	$([0.7, 5.2], 2)$	$(0.7, 5)$	—
Fig. 4 (c)	$(2.4, [0, 3])$	$(1.2, 3)$	$(0.7, 3.4)$
Fig. 4 (d)	$(2.4, 0.7)$	$(1.2, 3)$	$([0.7, 1.2], 3.4)$
Fig. 5	$(2.4, 0.7)$	$(1.2, 3)$	$(0.7, 3.4)$

where α_l and β_l correspond to the mean and standard deviation of the underlying normal distribution.

The mean μ_l and the variance σ_l^2 of $Y \sim Q_l$ are given by:

$$\mu_l = \exp(\alpha_l + \frac{\beta_l^2}{2}), l \in \mathcal{N}_{\text{Sat}} \tag{45a}$$

$$\sigma_l^2 = (\exp(\beta_l^2) - 1)\exp(2\alpha_l + \beta_l^2), l \in \mathcal{N}_{\text{Sat}}. \tag{45b}$$

6.2.2 Terrestrial Routes with Stochastic Delays. If $l \in \mathcal{N}_{\text{Ter}}$, we leverage the *gamma distribution* to simulate the statistics of delay y, where the probability density function is given by [16]:

$$P_{Y\sim Q_l}(y) = \frac{1}{\Gamma(\theta_l)\gamma_l^{\theta_l}} y^{\theta_l - 1} e^{-y/\gamma_l}, l \in \mathcal{N}_{\text{Ter}}. \tag{46}$$

The mean μ_l and the variance σ_l^2 of $Y \sim Q_l$ are given by:

$$\mu_l = \theta_l \gamma_l, \text{ and } \sigma_l^2 = \theta_l \gamma_l^2, l \in \mathcal{N}_{\text{Ter}}. \tag{47}$$

6.3 Parameter Settings

We consider a scenario with three available routes where $\mathcal{N} = \{1, 2, 3\}$. The parameter setting for the simulations is presented in

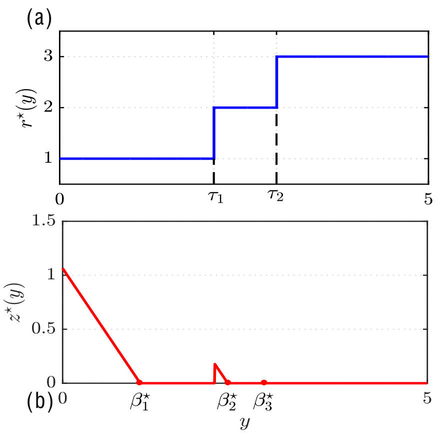

Figure 5: Visualization of simulated optimal policies.

Table 1. In this table, parameters specified as intervals indicate the values that are varied along the horizontal axis of the corresponding simulation figure. For notational convenience, we define:

$$\sigma_{\min} \triangleq \min_{r\in\mathcal{N}} \sigma_r, \quad \mu_{\min} \triangleq \min_{r\in\mathcal{N}} \mu_r, \quad \mu_{\max} \triangleq \max_{r\in\mathcal{N}} \mu_r. \tag{48}$$

6.4 Discussions

Fig. 4(a) highlights a surprising finding: *a higher delay variance may counterintuitively improve the average AoI performance.* In a single-route setting, reducing the delay variance typically leads to more regular update arrivals and thus lower AoI. However, this intuitive conclusion breaks down when an additional route is available, as evidenced by Fig. 4(a). Moreover, the figure shows that when $\sigma_{\min} = \sigma_1$ is below a certain threshold (approximately 2.5), the optimal policy actively utilizes both routes. Once this threshold is crossed, route 1 is no longer selected, and the route 2 (MAD-Optimal) with minimum delay is always used.

Fig. 4(b) demonstrates that route 1 (MDV-Optimal) provides better AoI performance than route 2 (MAD-Optimal) when $\mu_1 = \mu_{\max}$ is relatively small. While route 2 is used in the optimal policy due to its shorter average delay, its role is marginal. Notably, the benefit of joint routing peaks when $\mu_{\max}$ is just above 3. As $\mu_{\max}$ continues to increase and exceeds approximately 4, the route 2 (MAD-Optimal) policy becomes age-optimal.

Fig. 4(c) presents the long-term average AoI values in a three-route scenario, where $\sigma_1 = \sigma_{\min}$ is varied from 0 to 3. The optimal policy utilizes all three routes until $\sigma_{\min}$ exceeds a threshold of approximately 1.5, beyond which route 1 is no longer selected, and the policy relies solely on routes 2 and 3.

Fig. 4(d) shows the long-term average AoI values for another three-route scenario, where $\mu_3 = \mu_{\min}$ is varied from 0.7 to 1.2. In this case, all three routes consistently appear in the optimal policy, as route 3 remains the minimum delay route. However, its contribution to the overall performance becomes negligible as $\mu_3 = \mu_{\min}$ approaches $\mu_2 = 1.2$.

Fig. 5(a) and Fig. 5(b) show threshold structure of the optimal routing decision $R_i^\star = r^\star(y)$ and waiting time decision $Z_i^\star = z^\star(y)$

when $Y_i = y$. The parameter configuration is specified in Table 1. These results verify Theorem 1.

Overall, the proposed joint sampling and routing policy demonstrates robust improvements in AoI performance under diverse parameter settings. In particular, our simulations show that even in a basic two-route example, average AoI can be reduced by as much as 11%. This finding challenges conventional intuition and reveals a critical insight: *routes that appear suboptimal in isolation—due to higher mean delays or variances—can meaningfully contribute to AoI minimization under a well-designed optimized handover policy.*

7 Conclusion

In this work, we investigated a multi-route status update system and proved that a threshold-based joint sampling and routing policy can minimize the long-term average AoI. We introduced an efficient algorithm namely Bisec-ReaVI to compute this optimal policy. Our simulations consistently show improvements in AoI, revealing that higher variance or mean delays in certain routes can still help minimize AoI when jointly optimized. This challenges the common intuition that lower delay variance always leads to better AoI performance and provides insights into routing design for future TN-NTN networks.

Acknowledgments

This work is supported by the European Union through ERC Advanced Grant 101122990-GO SPACE-ERC-2023-A. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

[1] Baran Tan Bacinoglu, Elif Tugce Ceran, and Elif Uysal-Biyikoglu. 2015. Age of information under energy replenishment constraints. In IEEE ITA. 25–31.

[2] Baran Tan Bacinoglu, Yin Sun, Elif Uysal, and Volkan Mutlu. 2019. Optimal status updating with a finite-battery energy harvesting source. Journal of Communications and Networks 21, 3 (2019), 280–294.

[3] Ahmed M Bedewy, Yin Sun, Sastry Kompella, and Ness B Shroff. 2021. Optimal sampling and scheduling for timely status updates in multi-source networks. IEEE Transactions on Information Theory 67, 6 (2021), 4019–4034.

[4] Hasan Burhan Beytur and Elif Uysal-Biyikoglu. 2018. Minimizing age of information for multiple flows. In IEEE BlackSeaCom. IEEE, 1–5.

[5] Elif Tuğçe Ceran, Deniz Gündüz, and András György. 2021. A reinforcement learning approach to age of information in multi-user networks with HARQ. IEEE Journal on Selected Areas in Communications 39, 5 (2021), 1412–1426.

[6] Gongpu Chen, Soung Chang Liew, and Yulin Shao. 2022. Uncertainty-of-information scheduling: A restless multiarmed bandit framework. IEEE Transactions on Information Theory 68, 9 (2022), 6151–6173.

[7] Xiaomeng Chen, Aimin Li, and Shaohua Wu. 2024. Optimal Sampling for Uncertainty-of-Information Minimization in a Remote Monitoring System. In 2024 IEEE Information Theory Workshop (ITW). 115–120. doi:10.1109/ITW61385.2024.10807024

[8] Ronald A Howard. 1960. Dynamic programming and markov processes. Cambridge, MA; MIT Press.

[9] Igor Kadota, Abhishek Sinha, and Eytan Modiano. 2019. Scheduling algorithms for optimizing age of information in wireless networks with throughput constraints. IEEE/ACM Transactions on Networking 27, 4 (2019), 1359–1372.

[10] Igor Kadota, Abhishek Sinha, Elif Uysal-Biyikoglu, Rahul Singh, and Eytan Modiano. 2018. Scheduling policies for minimizing age of information in broadcast wireless networks. IEEE/ACM Transactions on Networking 26, 6 (2018), 2637–2650.

[11] Sanjit Kaul, Roy Yates, and Marco Gruteser. 2012. Real-time status: How often should one update?. In 2012 Proceedings IEEE INFOCOM. IEEE, 2731–2735.

[12] Shiyang Leng and Aylin Yener. 2019. Age of Information Minimization for an Energy Harvesting Cognitive Radio. IEEE Transactions on Cognitive Communications and Networking 5, 2 (2019), 427–439. doi:10.1109/TCCN.2019.2916097

[13] Aimin Li, Shaohua Wu, Gary C.F. Lee, Xiaomeng Chen, and Sumei Sun. 2024. Sampling to Achieve the Goal: An Age-aware Remote Markov Decision Process. In 2024 IEEE Information Theory Workshop (ITW). 121–126. doi:10.1109/ITW61385.2024.10806969

[14] Jiayu Pan, Ahmed M Bedewy, Yin Sun, and Ness B Shroff. 2021. Minimizing age of information via scheduling over heterogeneous channels. In ACM MobiHoc. 111–120.

[15] Jiayu Pan, Ahmed M. Bedewy, Yin Sun, and Ness B. Shroff. 2023. Age-Optimal Scheduling Over Hybrid Channels. IEEE Transactions on Mobile Computing 22, 12 (2023), 7027–7043. doi:10.1109/TMC.2022.3205292

[16] Peyton Z Peebles Jr. 2001. Probability, random variables, and random signal principles. McGraw-Hill.

[17] Sheldon M Ross. 1995. Stochastic processes. John Wiley & Sons.

[18] Md Kamran Chowdhury Shisher, Yin Sun, and I-Hong Hou. 2024. Timely communications for remote inference. IEEE/ACM Transactions on Networking (2024).

[19] Yin Sun and Benjamin Cyr. 2019. Sampling for data freshness optimization: Non-linear age functions. Journal of Communications and Networks 21, 3 (2019), 204–219.

[20] Yin Sun and Sastry Kompella. 2023. Age-optimal multi-flow status updating with errors: A sample-path approach. Journal of Communications and Networks 25, 5 (2023), 570–584.

[21] Yin Sun, Yury Polyanskiy, and Elif Uysal. 2019. Sampling of the Wiener process for remote estimation over a channel with random delay. IEEE Transactions on Information Theory 66, 2 (2019), 1118–1135.

[22] Yin Sun, Elif Uysal-Biyikoglu, Roy D Yates, C Emre Koksal, and Ness B Shroff. 2017. Update or wait: How to keep your data fresh. IEEE Transactions on Information Theory 63, 11 (2017), 7492–7508.

[23] Rajat Talak, Sertac Karaman, and Eytan Modiano. 2017. Minimizing age-of-information in multi-hop wireless networks. In 55th Annual Allerton Conference on Communication, Control, and Computing. IEEE, 486–493.

[24] Haoyue Tang, Yuchao Chen, Jintao Wang, Pengkun Yang, and Leandros Tassiulas. 2023. Age Optimal Sampling Under Unknown Delay Statistics. IEEE Transactions on Information Theory 69, 2 (2023), 1295–1314. doi:10.1109/TIT.2022.3233782

[25] Douglas J White. 1963. Dynamic programming, Markov chains, and the method of successive approximations. J. Math. Anal. Appl 6, 3 (1963), 373–376.

[26] Roy D Yates. 2015. Lazy is timely: Status updates by an energy harvesting source. In IEEE ISIT. 3008–3012.

[27] Abolfazl Zakeri, Mohammad Moltafet, and Marian Codreanu. 2024. Goal-oriented remote tracking of an unobservable multi-state Markov source. In 2024 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 1–6.

[28] Mengqiu Zhou, Meng Zhang, Howard H Yang, and Roy D Yates. 2024. Age-minimal CPU scheduling. In IEEE INFOCOM 2024-IEEE Conference on Computer Communications. IEEE, 401–410.

A Proof of Lemma 1

Part 1. We first prove that

$$\lambda^\star \le \lambda \iff h(\lambda) \le 0. \tag{49}$$

If $\lambda^\star \le \lambda$,

$$\exists\, \pi,\ \limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1} \mathbb{E}_\pi[q(Y_i,\, Z_i,\, Y_{i+1})]}{\sum_{i=0}^{n-1} \mathbb{E}_\pi[Z_i + Y_{i+1}]} \le \lambda. \tag{50}$$

Moving λ to the left-hand side yields: $\exists\, \pi$

$$\limsup_{n\to\infty} \frac{\frac{1}{n}\sum_{i=0}^{n-1}\Big(\mathbb{E}_\pi\big[q(Y_i,\, Z_i,\, Y_{i+1})\big] - \lambda\, \mathbb{E}_\pi[Z_i + Y_{i+1}]\Big)}{\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_\pi[Z_i + Y_{i+1}]} \le 0. \tag{51}$$

Since Y_i's over the same route are independent, the inter-sampling times $T_i = Y_i + Z_i$ are regenerative. Since there are N routes, the expected period of the most frequently used route satisfies $\mathbb{E}[n_{k+1} - n_k] \le N$, where n_k denotes the k-th epoch a particular route is used. Because T_i's are regenerative and we have $0 < \mathbb{E}[D_{n_{k+1}} - D_{n_k}] < \infty$, for all k, the renewal theory [17] tells us that $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[Z_i + Y_{i+1}]$ exists and is positive. Thus, there exists a policy π such that the numerator of the left-hand side (51) is less than zero, which indicates that the infimum of the numerator in (51) is less than zero, indicating that $h(\lambda) \le 0$.

Conversely, if $h(\lambda) \leq 0$, as $\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}[Z_i + Y_{i+1}]$ exists and is positive, we can derive (50) and (51), which indicates that $\lambda^\star \leq \lambda$. The corollary $\lambda^\star > \lambda \Longleftrightarrow h(\lambda) > 0$ can be derived directly from (49) by leveraging *Modus Tollens*.

Part 2: $\lambda^\star = \lambda \Longleftrightarrow h(\lambda) = 0$. If $h(\lambda) = 0$, from part 1 of the proof, we can first establish that $\lambda^\star \leq \lambda$. We then show that the policy π such that $h(\lambda) = 0$ can lead to

$$\limsup_{n\to\infty} \frac{\sum_{i=0}^{n-1}\mathbb{E}_\pi[q(Y_i,\ Z_i,\ Y_{i+1})]}{\sum_{i=0}^{n-1}\mathbb{E}_\pi[Z_i + Y_{i+1}]} = \lambda, \tag{52}$$

which indicates that $\lambda \geq \lambda^\star$. Combining these together, we can obtain $\lambda = \lambda^\star$. Conversely, if $\lambda = \lambda^\star$, we can establish from part 1 that $h(\lambda) \leq 0$; Meanwhile, the definition of $\lambda^\star$ in (5) leads to

$$\forall\ \pi,\ \limsup_{n\to\infty} \sum_{i=0}^{n-1}\mathbb{E}_\pi[q(Y_i,\ Z_i,\ Y_{i+1})] - \lambda\mathbb{E}_\pi[Z_i + Y_{i+1}] \geq 0, \tag{53}$$

which indicates that $h(\lambda) \geq 0$. Combining these together, we establish that $h(\lambda) = 0$.

B Proof of Lemma 5

Differentiating the action-value function given in (32) with respect to y, we obtain:

$$\frac{\partial Q(y,r)}{\partial y} = \begin{cases} \lambda^\star - y, & \text{if } y < \lambda^\star - \mu_r \\ \mu_r, & \text{if } y \geq \lambda^\star - \mu_r \end{cases} \tag{54}$$

For all y, the derivative is positive. Hence, $\forall r \in \mathcal{N}$, $Q(y,r)$ is monotonically increasing with y. As a result, for any $y_2 \geq y_1$, we can establish that

$$\begin{aligned} V^*(y_2;\lambda^\star) - V^*(y_1;\lambda^\star) &= \min_r Q(y_2,r) - \min_r Q(y_1,r) \\ &\geq \min_r\{Q(y_2,r) - Q(y_1,r)\} \geq 0, \end{aligned} \tag{55}$$

which indicates that $V^*(y;\lambda^\star) = \min_r Q(y,r)$ is monotonically increasing with y.

Since $\mu_j > \mu_k$, it follows that $\lambda^\star - \mu_k > \lambda^\star - \mu_j$. Then, using (54), we compute the difference:

$$\frac{\partial Q(y,j)}{\partial y} - \frac{\partial Q(y,k)}{\partial y} = \begin{cases} 0, & \text{if } y < \lambda^\star - \mu_j \\ \mu_j + y - \lambda^\star, & \text{if } \lambda^\star - \mu_k > y \geq \lambda^\star - \mu_j \\ \mu_j - \mu_k, & \text{if } y \geq \lambda^\star - \mu_k. \end{cases} \tag{56}$$

In all cases, the difference is non-negative, thus

$$\frac{\partial Q(y,j)}{\partial y} - \frac{\partial Q(y,k)}{\partial y} \geq 0, \tag{57}$$

which completes the proof.

C Proof of Lemma 7

As given by (54), the sole dependence of $\partial Q(y,i)/\partial y$ on i is μ_i. Thus, we have

$$\frac{\partial Q(y,j)}{\partial y} = \frac{\partial Q(y,k)}{\partial y}, \forall y \in \mathbb{R}^+, \forall j,k \in \mathcal{G}_\mu. \tag{58}$$

Thus, if route i is optimal at $y = 0$, it is also optimal for every $y \in \mathbb{R}^+$:

$$Q(y,i) = \min_{r\in\mathcal{G}_\mu}\{Q(y,r)\}, \forall y \in \mathbb{R}^*, \tag{59}$$

which indicates that this route *dominates* the space $\mathcal{G}_\mu$. As a result, only one route from $\mathcal{G}_\mu$ will be included in $\mathcal{R}^\star$.

D Proof of Lemma 2

Let $Q_2(\mu_2,\sigma_2)$ be any fixed delay distribution. Because $\mu_2 < \mu_1$, route 2 becomes optimal for sufficiently large values of y. One can construct a distribution $Q_1(\mu_1,\sigma_1)$ such that its age performance under the optimal waiting policy [22] satisfies:

$$\lambda_1 < \lambda_2, \tag{60}$$

where λ_j denotes the long-term average age achieved by exclusively using route j under the optimal waiting policy. Further, by the definition of optimality, the system-optimal age must satisfy:

$$\lambda^\star \leq \min(\lambda_1, \lambda_2). \tag{61}$$

Combining (60) and (61) we obtain:

$$\lambda^\star \leq \lambda_1 < \lambda_2. \tag{62}$$

This inequality implies that the optimal policy cannot exclusively use route 2; it must include route 1 as a part of the solution. Therefore, for distributions $Q_1(\mu_1,\sigma_1)$ satisfying (60), a threshold τ_1 must exist. This completes the proof.

E Proof of Theorem 2

As established from Appendices C and D, a route $k \in \mathcal{G}_{\mu_N}$ is optimal for sufficiently large values of y. Therefore, it must be used in the optimal policy and no route j, $j \notin \mathcal{G}_{\mu_N}$, policy can be optimal.

Leveraging Lemma 6, and setting $y^\star = 0$, we know that if route k is optimal at $y = 0$, it remains optimal $\forall y$, under the threshold policy structure. Therefore, if we have

$$Q(0,k) = \min_r Q(0,r) \Longleftrightarrow \arg\min_r Q(0,r) = k \tag{63}$$

then it is selected for all values of y, and the resulting policy is trivial. If, route k does not satisfy (63), another route will be present in the optimal solution. This completes the proof.

F Proof of Lemma 4

Consider a policy $\pi = (k, 0, k, 0, \ldots)$ that selects a single route k and generates a new sample immediately after the previous update packet is delivered (i.e., $Z_i = 0,\ \forall i$). We denote the long-term average age under this zero-wait route k policy as λ_k^{zw}. This age simplifies to $\frac{3\mu_k}{2} + \frac{\sigma_k^2}{2\mu_k}$ since Y_i's over a single route are i.i.d.

Let λ^π denote the average age achieved by a policy $\pi \in \Pi$. By definition of optimality, we have

$$\lambda^\star \leq \lambda^\pi, \quad \forall \pi \in \Pi. \tag{64}$$

In particular, this implies: $\lambda^\star \leq \min_r \lambda_r^{zw} = \min_r\{\frac{3\mu_r}{2} + \frac{\sigma_r^2}{2\mu_r}\}$. Since the AoI $\Delta(t)$ is non-negative, $\lambda^\star \geq 0$. This completes the proof.